

```
PI Slangenaupt S, Guseella JF;
XX
XX MPI; 2002-674806/72.
DR
XX
XX New IKKAP genes with mutations, useful for identifying a subject with
PT Familial dysautonomia (FD), or for rapid carrier screening in the
PT Ashkenazi Jewish population, e.g. screening presymptomatic homozygotes or
XX prenatal diagnosis.
XX
PS Claim 1; Page; 109pp; English.
XX
XX The present invention relates to methods and compositions useful for
CC detecting mutations which cause Familial Dysautonomia (FD, Riley-Day
CC syndrome, Hereditary Sensory and Autonomic Neuropathy Type III) (OMIM
CC 223900). It was found that mutations in the IKKAP gene (see AB080565)
CC are associated with FD. The mutation associated with the major haplotype
CC of FD, FDI mutation, is a base pair (bp) mutation, where the thymine
CC nucleotide located at bp 6 of intron 20 in the IKKAP gene is replaced
CC with a cytosine. This results in skipping of exon 20 in the mRNA from FD
CC patients, although they continue to express varying levels of wild-type
CC message in a tissue-specific manner. The mutation associated with the
CC minor haplotype, FD2 mutation, is a bp mutation, where the guanine
CC nucleotide at bp 2397 (bp 73 of exon 19) is replaced with a cytosine.
CC This bp mutation causes an arginine to proline missense mutation (R696P)
CC in the IKKAP protein, which is predicted to disrupt a potential
CC phosphorylation site. The IKKAP nucleic acid sequences are useful for
CC identifying a subject with FD and for rapid carrier screening. The IKKAP
CC gene contains 37 exons and maps to chromosome 9q31. Note: the present
CC sequence was not shown in the specification, but was derived from the
CC human wild-type IKKAP sequence given in Fig 6
XX
SQ Sequence 66479 BP; 18271 A; 12399 C; 14128 G; 21681 T; 0 U; 0 Other;
XX
Query Match 100.0%; Score 20; DB 6; Length 66479;
Best Local Similarity 100.0%; Pred. No. 5.1;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AGTGGCAACAGTACATGG 20
DB 34222 AGTGGCAACAGTACATGG 34203
XX
RESULT 3
AB080566/c
ID AB080566 standard; DNA; 66479 BP.
XX
XX AB080566;
AC
XX 08-NOV-2002 (first entry)
DT
XX
XX Mutant human IKKAP gene #1.
DE
XX
XX Human; IKKAP; Familial Dysautonomia; FD; Riley-Day syndrome;
KW Hereditary Sensory and Autonomic Neuropathy Type III; carrier screening;
XX FDI; mutation; gene; chromosome 9q31; ds.
XX
XX Homo sapiens.
OS Synthetic.
XX
XX Key Location/Qualifiers
FH mutation replace(34201,T)
FT /*tag= a
FT
XX
XX WO200259381-A2.
XX
XX 01-AUG-2002.
XX
XX 07-JAN-2002; 2002MO-US000473.
XX
XX 06-JAN-2001; 2001US-0260080P.
XX
XX (GEHO ) GEN HOSPITAL CORP.
XX
```

```
PI Slangenaupt S, Guseella JF;
XX
XX MPI; 2002-674806/72.
DR
XX
XX New IKKAP genes with mutations, useful for identifying a subject with
PT Familial dysautonomia (FD), or for rapid carrier screening in the
PT Ashkenazi Jewish population, e.g. screening presymptomatic homozygotes or
XX prenatal diagnosis.
XX
PS Claim 1; Page; 109pp; English.
XX
XX The present invention relates to methods and compositions useful for
CC detecting mutations which cause Familial Dysautonomia (FD, Riley-Day
CC syndrome, Hereditary Sensory and Autonomic Neuropathy Type III) (OMIM
CC 223900). It was found that mutations in the IKKAP gene (see AB080565)
CC are associated with FD. The mutation associated with the major haplotype
CC of FD, FDI mutation, is a base pair (bp) mutation, where the thymine
CC nucleotide located at bp 6 of intron 20 in the IKKAP gene is replaced
CC with a cytosine. This results in skipping of exon 20 in the mRNA from FD
CC patients, although they continue to express varying levels of wild-type
CC message in a tissue-specific manner. The mutation associated with the
CC minor haplotype, FD2 mutation, is a bp mutation, where the guanine
CC nucleotide at bp 2397 (bp 73 of exon 19) is replaced with a cytosine.
CC This bp mutation causes an arginine to proline missense mutation (R696P)
CC in the IKKAP protein, which is predicted to disrupt a potential
CC phosphorylation site. The IKKAP nucleic acid sequences are useful for
CC identifying a subject with FD and for rapid carrier screening. The IKKAP
CC gene contains 37 exons and maps to chromosome 9q31. Note: the present
CC sequence was not shown in the specification, but was derived from the
CC human wild-type IKKAP sequence given in Fig 6
XX
SQ Sequence 66479 BP; 18271 A; 12399 C; 14129 G; 21680 T; 0 U; 0 Other;
XX
Query Match 100.0%; Score 20; DB 6; Length 66479;
Best Local Similarity 100.0%; Pred. No. 5.1;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AGTGGCAACAGTACATGG 20
DB 34222 AGTGGCAACAGTACATGG 34203
XX
RESULT 4
AB080568/c
ID AB080568 standard; DNA; 66479 BP.
XX
XX AB080568;
AC
XX 08-NOV-2002 (first entry)
DT
XX
XX Mutant human IKKAP gene #3.
DE
XX
XX Human; IKKAP; Familial Dysautonomia; FD; Riley-Day syndrome;
KW Hereditary Sensory and Autonomic Neuropathy Type III; carrier screening;
XX FDI; FD2; mutation; gene; chromosome 9q31; ds.
XX
XX Homo sapiens.
OS Synthetic.
XX
XX Key Location/Qualifiers
FH mutation replace(33714,G)
FT /*tag= a
FT mutation replace(34201,T)
FT /*tag= b
FT
XX
XX WO200259381-A2.
XX
XX 01-AUG-2002.
XX
XX 07-JAN-2002; 2002MO-US000473.
XX
XX 06-JAN-2001; 2001US-0260080P.
XX
XX
```

PA (GEHO) GEN HOSPITAL CORP.
XX Staugenhaupt S, Gusella JF;
PI WPI; 2002-674806/72.
XX
PT New IKKAP genes with mutations, useful for identifying a subject with
PT familial dysautonomia (FD), or for rapid carrier screening in the
PT Ashkenazi Jewish population, e.g. screening presymptomatic homozygotes or
PT prenatal diagnosis.
XX
PS Claim 1; Page; 109pp; English.
XX
CC The present invention relates to methods and compositions useful for
CC detecting mutations which cause Familial Dysautonomia (FD, Riley-Day
CC syndrome, Hereditary Sensory and Autonomic Neuropathy Type III) (OMIM
CC 223900). It was found that mutations in the IKKAP gene (see AB080565)
CC are associated with FD. The mutation associated with the major haplotype
CC of FD, FDI mutation, is a base pair (bp) mutation, where the thymine
CC nucleotide located at bp 6 of intron 20 in the IKKAP gene is replaced
CC with a cytosine. This results in skipping of exon 20 in the mRNA from FD
CC patients, although they continue to express varying levels of wild-type
CC message in a tissue-specific manner. The mutation associated with the
CC minor haplotype, FD2 mutation, is a bp mutation, where the guanine
CC nucleotide at bp 2397 (bp 73 of exon 19) is replaced with a cytosine.
CC This bp mutation causes an arginine to proline missense mutation (R696P)
CC in the IKKAP protein, which is predicted to disrupt a potential
CC phosphorylation site. The IKKAP nucleic acid sequences are useful for
CC identifying a subject with FD and for rapid carrier screening. The IKKAP
CC gene contains 37 exons and maps to chromosome 9q31. Note: the present
CC sequence was not shown in the specification, but was derived from the
CC human wild-type IKKAP sequence given in Fig 6
XX
SQ Sequence 66479 BP; 18271 A; 12400 C; 14128 G; 21680 T; 0 U; 0 Other;
XX
Query Match 100.0%; Score 20; DB 6; Length 66479;
Best Local Similarity 100.0%; Pred. No. 5.1;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 AGTCGCAACAGTACATGG 20
Db 34222 AGTCGCAACAGTACATGG 34203
XX
RESULT 5
AB080565/c
ID AB080565 standard; DNA; 66479 BP.
XX
AC AB080565;
XX
DT 08-NOV-2002 (first entry)
XX
XX Human IKKAP wild-type gene.
XX
DE Human; IKKAP; Familial Dysautonomia; FD; Riley-Day syndrome;
KW Hereditary Sensory and Autonomic Neuropathy Type III; carrier screening;
KW gene; chromosome 9q31; ds.
XX
XX Homo sapiens.
XX
XX W0200259381-A2.
XX
XX 01-AUG-2002.
XX
XX 07-JAN-2002; 2002W0-US000473.
XX
XX 06-JAN-2001; 2001US-0260080P.
XX
XX (GEHO) GEN HOSPITAL CORP.
XX
XX Staugenhaupt S, Gusella JF;
XX
XX WPI; 2002-674806/72.
XX
XX

XX
XX New IKKAP genes with mutations, useful for identifying a subject with
XX familial dysautonomia (FD), or for rapid carrier screening in the
XX Ashkenazi Jewish population, e.g. screening presymptomatic homozygotes or
XX prenatal diagnosis.
XX
XX Claim 1; Fig 6; 109pp; English.
XX
XX The present invention relates to methods and compositions useful for
XX detecting mutations which cause Familial Dysautonomia (FD, Riley-Day
XX syndrome, Hereditary Sensory and Autonomic Neuropathy Type III) (OMIM
XX 223900). It was found that mutations in the IKKAP gene (the present
XX sequence) are associated with FD. The mutation associated with the major
XX haplotype of FD, FDI mutation, is a base pair (bp) mutation, where the
XX thymine nucleotide located at bp 6 of intron 20 in the IKKAP gene is
XX replaced with a cytosine. This results in skipping of exon 20 in the mRNA
XX from FD patients, although they continue to express varying levels of
XX wild-type message in a tissue-specific manner. The mutation associated
XX with the minor haplotype, FD2 mutation, is a bp mutation, where the
XX guanine nucleotide at bp 2397 (bp 73 of exon 19) is replaced with a
XX cytosine. This bp mutation causes an arginine to proline missense
XX mutation (R696P) in the IKKAP protein, which is predicted to disrupt a
XX potential phosphorylation site. The IKKAP nucleic acid sequences are
XX useful for identifying a subject with FD and for rapid carrier screening.
XX The IKKAP gene contains 37 exons and maps to chromosome 9q31
XX
XX Sequence 66479 BP; 18271 A; 12398 C; 14128 G; 21681 T; 0 U; 0 Other;
XX
XX
XX Query Match 100.0%; Score 20; DB 6; Length 66479;
XX Best Local Similarity 100.0%; Pred. No. 5.1;
XX Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX
XX 1 AGTCGCAACAGTACATGG 20
XX Db 34222 AGTCGCAACAGTACATGG 34203
XX
XX
XX RESULT 6
XX ACF67367_52/c (53 of 57) of ACF67367 from base 5200001 (Photorehabus luminescens nucleotic
XX Continuation split into 57 fragments LOCUS ACF67367 Accession ACF67367
XX WP Sequence Fragment Name Begin End
XX WP ACF67367_00 1 110000
XX WP ACF67367_01 100001 210000
XX WP ACF67367_02 200001 310000
XX WP ACF67367_03 300001 410000
XX WP ACF67367_04 400001 510000
XX WP ACF67367_05 500001 610000
XX WP ACF67367_06 600001 710000
XX WP ACF67367_07 700001 810000
XX WP ACF67367_08 800001 910000
XX WP ACF67367_09 900001 1010000
XX WP ACF67367_10 1000001 1110000
XX WP ACF67367_11 1100001 1210000
XX WP ACF67367_12 1200001 1310000
XX WP ACF67367_13 1300001 1410000
XX WP ACF67367_14 1400001 1510000
XX WP ACF67367_15 1500001 1610000
XX WP ACF67367_16 1600001 1710000
XX WP ACF67367_17 1700001 1810000
XX WP ACF67367_18 1800001 1910000
XX WP ACF67367_19 1900001 2010000
XX WP ACF67367_20 2000001 2110000
XX WP ACF67367_21 2100001 2210000
XX WP ACF67367_22 2200001 2310000
XX WP ACF67367_23 2300001 2410000
XX WP ACF67367_24 2400001 2510000
XX WP ACF67367_25 2500001 2610000
XX WP ACF67367_26 2600001 2710000
XX WP ACF67367_27 2700001 2810000
XX WP ACF67367_28 2800001 2910000
XX WP ACF67367_29 2900001 3010000
XX WP ACF67367_30 3000001 3110000
XX WP ACF67367_31 3100001 3210000

WP ACF67367_31 3100001 3210000
 WP ACF67367_32 3200001 3310000
 WP ACF67367_33 3300001 3410000
 WP ACF67367_34 3400001 3510000
 WP ACF67367_35 3500001 3610000
 WP ACF67367_36 3600001 3710000
 WP ACF67367_37 3700001 3810000
 WP ACF67367_38 3800001 3910000
 WP ACF67367_39 3900001 4010000
 WP ACF67367_40 4000001 4110000
 WP ACF67367_41 4100001 4210000
 WP ACF67367_42 4200001 4310000
 WP ACF67367_43 4300001 4410000
 WP ACF67367_44 4400001 4510000
 WP ACF67367_45 4500001 4610000
 WP ACF67367_46 4600001 4710000
 WP ACF67367_47 4700001 4810000
 WP ACF67367_48 4800001 4910000
 WP ACF67367_49 4900001 5010000
 WP ACF67367_50 5000001 5110000
 WP ACF67367_51 5100001 5210000
 WP ACF67367_52 5200001 5310000
 WP ACF67367_53 5300001 5410000
 WP ACF67367_54 5400001 5510000
 WP ACF67367_55 5500001 5610000
 WP ACF67367_56 5600001 5648894

Query Match Best Local Similarity 84.0%; Score 16.8; DB 10; Length 110000;
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AGTGCACAAACAGTACAAATGG 20

Db 43881 AGACGCAAAACAGTACAAAG 43862

RESULT 7

Continuation (4 of 7) of ACF65387 from base 300001 (Photographus luminescens nucleotide
 WP Sequence split into 7 fragments LOCUS ACF65387 Accession ACF65387

WP	Fragment Name	Begin	End
WP ACF65387_0		1	110000
WP ACF65387_1		100001	210000
WP ACF65387_2		200001	310000
WP ACF65387_3		300001	410000
WP ACF65387_4		400001	510000
WP ACF65387_5		500001	610000
WP ACF65387_6		600001	696798

Query Match Best Local Similarity 84.0%; Score 16.8; DB 10; Length 110000;
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AGTGCACAAACAGTACAAATGG 20

Db 83395 AGACGCAAAACAGTACAAAG 83376

RESULT 8

ABL03413 ABL03413 standard; cDNA; 950 BP.

AC ABL03413;

XX 26-MAR-2002 (first entry)

XX Drosophila melanogaster expressed polynucleotide SEQ ID NO 4721.

XX Drosophila; developmental biology; cell signalling; insecticide;

XX pharmaceutical; gene; ss.

XX Drosophila melanogaster.

PN WO200171042-A2.

XX 27-SEP-2001.

XX 23-MAR-2001; 2001WO-US009231.

XX 23-MAR-2000; 2000US-0191637P.

XX 11-JUL-2000; 2000US-00614150.

XX (PEKE) PE CORP NY.

XX Venter JC, Adams M, Li PWD, Myers EW;

XX WPI; 2001-656860/75.

XX P-PSDB; ABB59310.

XX New isolated nucleic acid detection reagent for detecting 1000 or more

XX genes from Drosophila and for elucidating cell signaling and cell-cell

XX interactions.

XX Claim 1; SEQ ID NO 4721; 21pp + Sequence Listing; English.

XX The invention relates to an isolated nucleic acid detection reagent

XX capable of detecting 1000 or more genes from Drosophila. The invention is

XX useful in developmental biology and in elucidating cell signalling and

XX cell-cell interactions in higher eukaryotes for the development of

XX insecticides, therapeutics and pharmaceutical drugs. The invention

XX discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA

XX sequences (AB01840-ABL16175) and the encoded proteins (ABB57737-

XX ABB32072). The sequence data for this patent did not form part of the

XX printed specification, but was obtained in electronic format directly

XX from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX Sequence 950 BP; 229 A; 234 C; 248 G; 239 T; 0 U; 0 Other;

XX Query Match Best Local Similarity 82.0%; Score 16.4; DB 4; Length 950;

XX Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

XX QY 3 TCGCAACAGTACAAATGG 20

XX Db 634 TCGCAACAGTACAAATGG 651

XX RESULT 9

XX ABL16514/c

XX ID ABL16514 standard; DNA; 2901 BP.

XX ABL16514;

XX 26-MAR-2002 (first entry)

XX Drosophila melanogaster genomic polynucleotide SEQ ID NO 1015.

XX Drosophila; developmental biology; cell signalling; insecticide;

XX pharmaceutical; gene; ds.

XX Drosophila melanogaster.

XX New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from *Drosophila* and for elucidating cell signaling and cell-cell
PT interactions.
XX
PS Claim 1; SEQ ID NO 1015; 21bp + Sequence Listing; English.
XX
CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from *Drosophila*. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (AB116176-AB130511), expressed DNA
CC sequences (AB101840-AB116175) and the encoded proteins (AB57737-
CC AB872072). The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pcr_sequences
XX
SQ Sequence 2901 BP; 817 A; 687 C; 668 G; 729 T; 0 U; 0 Other;
XX
Query Match 82.0%; Score 16.4; DB 4; Length 2901;
Best Local Similarity 94.4%; Pred. No. 2.5e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
QY 3 TCGCAACAGTACATGG 20
DB 2408 TCGCCACAGTACATGG 2391
XX
RESULT 10
AB103412/C
ID ABL03412 standard; cDNA; 3165 BP.
XX
XX ABL03412;
XX
DT 26-MAR-2002 (first entry)
XX
DE *Drosophila melanogaster* expressed polynucleotide SEQ ID NO 4718.
XX
XX *Drosophila*; developmental biology; cell signalling; insecticide;
XX pharmaceutical; gene; ss.
XX
OS *Drosophila melanogaster*.
XX
FN WO200171042-A2.
XX
XX 27-SEP-2001.
XX
PD 23-MAR-2001; 2001WO-US009231.
XX
PE 23-MAR-2001; 2000US-0191637P.
XX
PR 23-MAR-2000; 2000US-0191637P.
XX
PR 11-JUL-2000; 2000US-00614150.
XX
PA (PEKE) PE CORP NY.
XX
PI Venter JC, Adams M, Li FWD, Myers EW,
XX
DR WPI; 2001-656860/75.
XX
DR P-PSDB; ABB59309.
XX
XX New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from *Drosophila* and for elucidating cell signaling and cell-cell
PT interactions.
XX
PS Claim 1; SEQ ID NO 4718; 21bp + Sequence Listing; English.
XX
CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from *Drosophila*. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (AB116176-AB130511), expressed DNA
CC sequences (AB101840-AB116175) and the encoded proteins (AB57737-

CC AB872072). The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pcr_sequences
XX
SQ Sequence 3165 BP; 866 A; 763 C; 700 G; 836 T; 0 U; 0 Other;
XX
Query Match 82.0%; Score 16.4; DB 4; Length 3165;
Best Local Similarity 94.4%; Pred. No. 2.5e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
QY 3 TCGCAACAGTACATGG 20
DB 1317 TCGCCACAGTACATGG 1300
XX
RESULT 11
AAK54574
ID AAK54574 standard; cDNA; 182 BP.
XX
AC AAK54574;
XX
DT 13-NOV-2001 (first entry)
XX
DE Human haematological malignancy-related antigen coding sequence #299.
XX
XX Human; cytostatic; vascular; gene therapy; vaccine; lymphoma;
XX haematological malignancy; antigen; chronic lymphocytic leukaemia;
XX follicular lymphoma; Hodgkin's lymphoma; non-Hodgkin's lymphoma; ss.
XX
OS Homo sapiens.
XX
FN WO200164886-A2.
XX
XX 07-SEP-2001.
XX
PD 01-MAR-2001; 2001WO-US007272.
XX
PE 01-MAR-2001; 2000US-0186126P.
XX
PR 17-MAR-2000; 2000US-0190478P.
XX
PR 27-APR-2000; 2000US-0200545P.
XX
PR 28-APR-2000; 2000US-0200303P.
XX
PR 28-APR-2000; 2000US-0200779P.
XX
PR 01-MAY-2000; 2000US-0200999P.
XX
PR 04-MAY-2000; 2000US-0202084P.
XX
PR 22-MAY-2000; 2000US-0206201P.
XX
PR 14-JUL-2000; 2000US-0218958P.
XX
PR 03-AUG-2000; 2000US-0222903P.
XX
PR 04-AUG-2000; 2000US-0223416P.
XX
PR 07-AUG-2000; 2000US-0223378P.
XX
PA (CORI) CORIXA CORP.
XX
XX Gaiger A, Algate PA, Mannion J;
XX
XX WPI; 2001-514842/56.
XX
XX Compositions and methods for the detection of hematological malignancies,
PT e.g. chronic lymphocytic leukemia, lymphoma, follicular lymphoma and
PT Hodgkin's and T/B cell non-Hodgkin's lymphoma.
XX
XX Claim 31; Page 411; 1252pp; English.
XX
PS The present invention relates to compositions and methods for the
CC detection, diagnosis and therapy of haematological malignancies. The
CC present sequence is the coding sequence of a human haematological
CC malignancy related antigen. The methods of the present invention comprise
CC detecting the presence of haematological malignancy related antigen(s) in
CC a sample obtained from the patient (an increased level of the
CC polypeptide, compared to an unaffected individual, is indicative of an
CC increased risk). Haematological malignancies which can be treated using
CC the present invention are chronic lymphocytic leukaemia, lymphoma,
CC follicular lymphoma, Hodgkin's lymphoma, non-Hodgkin's lymphoma and T/B
CC cell non-Hodgkin's lymphoma

XX SQ Sequence 182 BP; 49 A; 38 C; 48 G; 45 T; 0 U; 2 Other;
Query Match 79.0%; Score 15.8; DB 4; Length 182;
Best Local Similarity 89.5%; Pred. No. 3.6e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
CY 1 AGTCGCAACAGTACATG 19
100 AGTCGCAACAGTACATG 118
Db 100 AGTCGCAACAGTACATG 118
RESULT 12
AAK54728
ID AAK54728 standard; cDNA; 182 BP.
XX AC AAK54728;
XX DT 13-NOV-2001 (first entry)
XX DE Human haematological malignancy-related antigen coding sequence #453.
XX DE Human; cytostatic; vascular; gene therapy; vaccine; lymphoma;
XX KW haematological malignancy; antigen; chronic lymphocytic leukaemia;
XX KW follicular lymphoma; Hodgkin's lymphoma; non-Hodgkin's lymphoma; ss.
XX OS Homo sapiens.
XX PN WC200164886-A2.
XX PD 07-SEP-2001.
XX PF 01-MAR-2001; 2001WC-US007272.
XX PR 01-MAR-2000; 2000US-0186126P.
XX PR 17-MAR-2000; 2000US-0190479P.
XX PR 27-APR-2000; 2000US-0200545P.
XX PR 28-APR-2000; 2000US-0200303P.
XX PR 28-APR-2000; 2000US-0200779P.
XX PR 01-MAY-2000; 2000US-0200999P.
XX PR 04-MAY-2000; 2000US-0202084P.
XX PR 22-MAY-2000; 2000US-0206201P.
XX PR 14-JUL-2000; 2000US-0218950P.
XX PR 03-AUG-2000; 2000US-0222903P.
XX PR 04-AUG-2000; 2000US-0223416P.
XX PR 07-AUG-2000; 2000US-0223378P.
XX PA (CORI-) CORIXA CORP.
XX PI Gaiger A, Algate PA, Mannion J;
XX DR WPI; 2001-514842/56.
XX PT Compositions and methods for the detection of hematological malignancies,
XX PT e.g. chronic lymphocytic leukemia, lymphoma, follicular lymphoma and
XX PT Hodgkin's and T/B cell non-Hodgkin's lymphoma.
XX PS Claim 31; Page 451; 1252pp; English.
XX CC The present invention relates to compositions and methods for the
XX CC detection, diagnosis and therapy of haematological malignancies. The
XX CC present sequence is the coding sequence of a human haematological
XX CC malignancy related antigen. The methods of the present invention comprise
XX CC detecting the presence of haematological malignancy related antigen(s) in
XX CC a sample obtained from the patient (an increased level of the
XX CC polypeptide, compared to an unaffected individual, is indicative of an
XX CC increased risk). Haematological malignancies which can be treated using
XX CC the present invention are chronic lymphocytic leukaemia, lymphoma,
XX CC follicular lymphoma, Hodgkin's lymphoma, non-Hodgkin's lymphoma and T/B
XX CC cell non-Hodgkin's lymphoma
XX SC Sequence 182 BP; 49 A; 38 C; 48 G; 45 T; 0 U; 2 Other;

Query Match 79.0%; Score 15.8; DB 4; Length 182;
Best Local Similarity 89.5%; Pred. No. 3.6e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
CY 1 AGTCGCAACAGTACATG 19
100 AGTCGCAACAGTACATG 118
Db 100 AGTCGCAACAGTACATG 118
RESULT 13
ADF79775/c
ID ADF79775 standard; DNA; 431 BP.
XX AC ADF79775;
XX DT 26-FEB-2004 (first entry)
XX DE Leukaemia-related DNA sequence #331.
XX DE Cytostatic; Gene therapy; leukaemia; ss.
XX OS Unidentified.
XX PN WC2003039443-A2.
XX PD 15-MAY-2003.
XX PF 04-NOV-2002; 2002WC-EP012303.
XX PR 05-NOV-2001; 2001EP-00126244.
XX PR 30-APR-2002; 2002EP-00009758.
XX PA (DEK-) DEUT KREBSFORSCHUNGSZENTRUM.
XX PA (UYLU-) UNIV LUDWIG MAXIMILIANS.
XX PA (HAE/) HAEFELACH T.
XX PA (SCHC/) SCHROCH C.
XX PA (KERN/) KERN W.
XX PI Haefelach T, Schoch C, Kern W, Kohlmann A, Schmittger S, Dugas M;
XX PI Elis R, Broers B, Mergenthaler S,
XX DR WPI; 2003-505037/47.
XX PT Determining the subtype of leukemia cells and whether a patient sample
XX PT contains leukemia cells or other cells, useful for treating leukemia,
XX PT comprises determining the expression profile of a group of markers in a
XX PT patient sample.
XX PS Disclosure; SEQ ID NO 331; 2938pp; English.
XX CC The present invention relates to a method (M1) for determining the
XX CC subtype of leukemia cells and whether a patient sample contains
XX CC leukemia cells. The method comprises determining the expression profile
XX CC of a group of markers in a patient sample. The method is useful for
XX CC determining the presence of leukemia cells, its types or subtypes, and
XX CC for the preparation of a medicament for treating leukaemia.
XX SC Sequence 431 BP; 135 A; 104 C; 89 G; 103 T; 0 U; 0 Other;
Query Match 79.0%; Score 15.8; DB 10; Length 431;
Best Local Similarity 89.5%; Pred. No. 4e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
CY 1 AGTCGCAACAGTACATG 19
281 AGTCGCAACAGTACATG 263
Db 281 AGTCGCAACAGTACATG 263
RESULT 14
ADF81342/c
ID ADF81342 standard; DNA; 440 BP.
XX AC ADF81342;

DT 11-FEB-2003 (first entry)
XX S. pneumoniae type 4 strain coding region #881.
XX
XX Gene; ds: bacterial meningitis; pneumonia; sepsis; otitis media;
KM ear infection; antiinflammatory; antibacterial; immunostimulant;
XX auditory; respiratory; gene therapy; vaccine.
XX
XX Streptococcus pneumoniae; type 4 strain.
OS
XX WO200277021-A2.
PN
XX 03-OCT-2002.
PD
XX 27-MAR-2002; 2002W0-IB002163.
PF
XX 27-MAR-2001; 2001GB-00007658.
PR
XX (CHIR-) CHIRON SPA.
PA (GENO-) INST GENOMIC RES.
XX
XX Masignani V, Tettelin H, Fraser C;
PI
XX WPI: 2003-040579/03.
DR P-PSDB; ABU01306.
DR
XX New proteins and nucleic acid molecules from Streptococcus pneumoniae,
PT useful as medicaments for treating or preventing a disease or infection
PT due to streptococcus bacteria, such as pneumonia, sepsis, otitis media or
PT ear infection.
XX
XX Claim 6; SEQ ID NO 1761; 56pp; English.
PS
XX The invention relates to a protein comprising or having at least 50%
XX identity to any of the 2469 amino acid sequences, identified in the
CC specification (available on a computer readable format), or its fragment,
CC expressed from 2469 of 2489 identified DNA coding regions from the
CC Streptococcus pneumoniae type 4 strain genomic sequence appearing as
CC AB556454. Also included are an antibody which binds one of the proteins,
CC treating a patient by administering the protein, DNA or antibody (in a
CC composition), a kit comprising first and second primers, which are the
CC nucleic acid cited above or fragments between nucleotides 8-100 of a
CC sequence not defined in the specification, for amplifying a target
CC sequence contained within a Streptococcus nucleic acid sequence, where
CC the first primer is substantially complementary to the target sequence
CC and the second primer is substantially complementary to the complement of
CC the target sequence, and where the parts of the primers having
CC substantial complementarity define the termini of the target sequence to
CC be amplified, assay comprising contacting a test compound with the
CC protein, and determining whether the test compound binds to the protein
CC and a Streptococcus pneumoniae bacterium, where one or more genes
CC encoding the proteins has been rendered inactive. The proteins, nucleic
CC acid molecules, antibody and compositions are useful as medicaments for
CC treating or preventing a disease or infection due to streptococcus
CC bacteria, particularly S. pneumoniae, such as pneumonia, sepsis, otitis
CC media or ear infection. They are also useful in developing vaccines,
CC diagnostics and antibiotics. The methods are useful for identifying
CC immunodominant proteins. The present sequence is one of the 2469
CC identified coding regions from the genomic sequence. Note: The sequence
CC data for this patent did not form part of the printed specification, but
CC was obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences. (Updated on 27-OCT-2003 to
CC standardise OS field)
XX
XX Sequence 903 BP; 190 A; 205 C; 170 G; 338 T; 0 U; 0 Other;
SQ

Query Match 79.0%; Score 15.8; DB 10; Length 903;
Best Local Similarity 89.5%; Pred. No. 4.3e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

2 GTGCGAAACAGTACATGG 20
|||||
335 GTTCGAAAAAGAACATGG 317

RESULT 17
ADFO0300
ID ADFO0300 standard; DNA; 1281 BP.
XX
XX
XX ADFO0300;
AC
XX 12-FEB-2004 (first entry)
DT
XX
XX Bacterial polynucleotide #585.
DE
XX
XX Proteus mirabilis infection; bacterial infection; antibacterial;
KM immunostimulant; gene; ds.
XX
XX Proteus mirabilis.
OS
XX US6605709-B1.
PN
XX 12-AUG-2003.
PD
XX
XX 05-APR-2000; 2000US-00543681.
PF
XX
XX 09-APR-1999; 99US-0128706P.
PR
XX
XX (GENO-) GENOME THERAPEUTICS CORP.
PA
XX Breton GL;
PI
XX WPI: 2003-895291/82.
DR P-PSDB; ADF04472.
DR
XX New Proteus mirabilis polypeptides and polynucleotides, useful as
PT reagents for diagnosis of bacterial disease, as components of
PT antibacterial vaccines, as targets for antibacterial drugs, or as
PT biocontrol agents for plants.
XX
XX Disclosure; SEQ ID NO 585; 870pp; English.
PS
XX The invention relates to new Proteus mirabilis polypeptides and
XX polynucleotides. The invention also relates to antibodies against the
XX polypeptides, methods for producing the polypeptides, a method of
XX generating vaccines for immunising an individual against P. mirabilis, a
XX method for evaluating a compound for the ability to bind a P. mirabilis
XX polypeptide and a method for screening test compounds for anti-bacterial
XX activity. The polypeptides and polynucleotides are useful as molecular
XX targets for diagnosing, preventing and treating pathological conditions
XX resulting from bacterial infection, as reagents for diagnosis of
XX bacterial diseases, as components of antibacterial vaccines, as targets
XX for antibacterial drugs or as bio-control agents for plants. This
XX sequence represents a Proteus mirabilis polynucleotide of the invention.
SQ

Query Match 79.0%; Score 15.8; DB 10; Length 1281;
Best Local Similarity 89.5%; Pred. No. 4.5e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

2 GTGCGAAACAGTACATGG 20
|||||
315 GTTCGAAACATTCATATGG 333

RESULT 18
ACA21176
ID ACA21176 standard; DNA; 1674 BP.
XX
XX
XX ACA21176;
AC
XX 19-JUN-2003 (first entry)
DT
XX
XX Prokaryotic essential gene #2833.
DE
XX

KW Antisense; ds; prokaryotic essential gene; cell proliferation;
KM drug design; gene.
XX Acinetobacter baumannii.
OS
XX
XX W020027183-A2.
XX
XX 03-OCT-2002.
PD
XX
XX 21-MAR-2002; 2002MO-US009107.
PF
XX
XX 21-MAR-2001; 2001US-00815242.
PR 06-SEP-2001; 2001US-00948993.
PR 25-OCT-2001; 2001US-0342923P.
PR 08-FEB-2002; 2002US-00072851.
PR 06-MAR-2002; 2002US-0362699P.
XX
XX (ELITRA PHARM INC.
XX
XX Wang L, Zamudio C, Malone C, Haselback R, Ohlsen KL, Zyskind JW;
PI Wall D, Trawick JD, Carr GT, Yamamoto R, Forsyth RA, Xu HH;
PI WPI; 2003-029926/02.
XX P-PSDB; ABU17306.
DR
XX
XX New antisense nucleic acids, useful for identifying proteins or screening
PT for homologous nucleic acids required for cellular proliferation to
PT isolate candidate molecules for rational drug discovery programs.
XX
XX Claim 14; SEQ ID NO 9046; 1766pp; English.
PS
XX
XX The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence data for this patent
CC not form part of the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 1674 BP; 476 A; 340 C; 410 G; 448 T; 0 U; 0 Other;
Query Match 79.0%; Score 15.8; DB 8; Length 1674;
Best Local Similarity 89.5%; Pred. No. 4.7e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GTGGCAACAGTACATGG 20
DB 473 GTGGCAACATTCATATGG 491

RESULT 19
ADA29603
ID ADA29603 standard; DNA; 1704 BP.
XX
XX
XX ADA29603;
XX
XX 20-NOV-2003 (first entry)
DT
XX
XX DNA encoding Acinetobacter baumannii protein #890.
DE
XX
XX ds; gene; Acinetobacter baumannii; bacterial disease; antibacterial;
KM vaccine; plant biocontrol agent.
XX
XX Acinetobacter baumannii.
OS
XX
XX US6562958-B1.
PN
XX
XX 13-MAY-2003.
PD
XX
XX 04-JUN-1999; 99US-00328352.
PF
XX
XX 09-JUN-1998; 98US-0088701P.
PR
XX
XX (GENO-) GENOME THERAPEUTICS CORP.
PA
XX
XX Breton G, Bush D;
PI WPI; 2003-576092/54.
PI P-PSDB; ADA33729.
DR
XX
XX New Acinetobacter baumannii proteins and nucleic acids, useful as reagents
PT for diagnosing a bacterial disease, as components of antibacterial
PT vaccines, as targets for antibacterial drugs, or as biocontrol agents for
PT plants.
XX
XX Example; SEQ ID NO 890; 328pp; English.
PS
XX
XX The invention relates to isolated Acinetobacter baumannii nucleic acids.
CC The A. baumannii nucleic acids and polypeptides are useful as reagents
CC for diagnosing a bacterial disease, as components of antibacterial
CC vaccines, as targets for antibacterial drugs, to detect the presence of
CC A. baumannii and other Acinetobacter species in a sample, in screening
CC compounds for the ability to interfere with the A. baumannii life cycle
CC or to inhibit A. baumannii infection, and as biocontrol agents for
CC plants. The present sequence represents DNA encoding an A. baumannii
CC protein.
XX
SQ Sequence 1704 BP; 488 A; 342 C; 418 G; 456 T; 0 U; 0 Other;
Query Match 79.0%; Score 15.8; DB 9; Length 1704;
Best Local Similarity 89.5%; Pred. No. 4.7e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GTGGCAACAGTACATGG 20
DB 500 GTGGCAACATTCATATGG 518

RESULT 20
ABK84220/C
ID ABK84220 standard; cDNA; 2417 BP.
XX
XX
XX ABK84220;
XX
XX 14-AUG-2002 (first entry)
DT
XX
XX Human cDNA differentially expressed in granulocytic cells #791.
DE
XX
XX Human; ss; granulocytic cell; DNA chip; bacterial infection;
KM viral infection; parasitic infection; protozoal infection;
KM fungal infection; sterile inflammatory disease; psoriasis;
KM rheumatoid arthritis; glomerulonephritis; asthma; thrombosis;

KM cardiac reperfusion injury; renal reperfusion injury; ARDS;
 KM adult respiratory distress syndrome; inflammatory bowel disease;
 KM Crohn's disease; ulcerative colitis; periodontal disease;
 KM granulocyte activation; chronic inflammation; allergy.
 XX Homo sapiens.
 XX MO200228999-A2.
 XX 11-APR-2002.
 XX 03-OCT-2001; 2001MO-US030821.
 XX 03-OCT-2000; 2000US-0237189P.
 XX (GENE-) GENE LOGIC INC.
 XX Beazer-Barclay Y, Weissman SM, Yamaga S, Vockley J;
 XX WPI; 2002-435328/46.
 XX
 XX Detecting granulocyte activation by detecting differential expression of
 PT genes associated with granulocyte activation, which serves as diagnostic
 PT markers that is useful for monitoring disease states and drug toxicity.
 PS Claim 1; SEQ ID NO 791; 114pp; English.
 XX
 XX The invention relates to detecting (M1) granulocyte (GC) activation
 CC (GCA), by detecting the level of expression of gene(s) (Gs) identified by
 CC DNA chip analysis as given in the specification, and comparing the
 CC expression level to an expression level in an unactivated GC, where
 CC differential expression of Gs is indicative of GCA. Also included are
 CC modulating (M2) Gs by contacting GC with an agent that alters the
 CC expression of at least one gene in Gs; (2) screening (M3) for an agent
 CC capable of modulating GCA or an inflammation (especially chronic) in a
 CC tissue, an allergic response in a subject, exposure of a subject to a
 CC pathogen or sterile inflammatory disease using the gene expression
 CC profile; (3) detecting (M4) an inflammation (especially chronic) in a
 CC tissue, an allergic response in a subject, exposure of a subject to a
 CC pathogen or sterile inflammatory disease, by detecting the level of
 CC expression in a sample of the tissue of gene(s) from Gs, where the level
 CC of expression of the gene is indicative of inflammation; (4) treating
 CC (M5) an inflammation (especially chronic) or in a tissue, an allergic
 CC response in a subject, exposure of a subject to a pathogen or sterile
 CC inflammatory disease, by contacting a tissue having inflammation with an
 CC agent that modulates the expression of gene(s) from Gs in the tissue. M1
 CC is useful for detecting GCA; M2 is useful for modulating GCA; M3 is useful
 CC for screening an agent capable of modulating GCA preferably in an
 CC inflammation in a tissue; M4 is useful for detecting an inflammation
 CC (especially chronic) in a tissue, an allergic response in a subject,
 CC exposure of a subject to a pathogen or sterile inflammatory disease (e.g.
 CC psoriasis, rheumatoid arthritis, glomerulonephritis, asthma, thrombosis,
 CC cardiac reperfusion injury, renal reperfusion injury, ARDS, adult
 CC respiratory distress syndrome, inflammatory bowel disease, Crohn's
 CC disease, ulcerative colitis, periodontal disease; also bacterial
 CC infection, viral infection, parasitic infection, protozoal infection,
 CC fungal infection and MS is useful for treating one of the above
 CC conditions. The present sequence represents a gene differentially
 CC expressed in granulocytes. Note: The sequence data for this patent did
 CC not form part of the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 XX Sequence 2417 BP; 788 A; 503 C; 521 G; 605 T; 0 U; 0 Other;
 SQ
 Query Match 79.0%; Score 15.8; DB 6; Length 2417;
 Best Local Similarity 89.5%; Pred. No. 4.9e+02;
 Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

1 AGTCGCAACAGTACATG 19
 |||||
 DB 2101 AGTCGCAACAGTACATG 2083

RESULT 21
 ADO19070/c
 ID ADO19070 standard; cDNA; 2417 BP.
 XX
 XX ADO19070;
 XX
 XX 12-AUG-2004 (first entry)
 XX
 XX Human PRO polynucleotide #4.
 XX
 XX Human, PRO, gene; ss; immune related disorder;
 KM systemic lupus erythematosus; rheumatoid arthritis; osteoarthritis;
 KM juvenile chronic arthritis; systemic sclerosis; Sjogren's syndrome;
 KM vasculitis; sarcoidosis; autoimmune haemolytic anaemia;
 KM autoimmune thrombocytopenia; thyroiditis; diabetes mellitus;
 KM renal disease; demyelinating disease; central nervous system;
 KM peripheral nervous system; demyelinating polyneuropathy;
 KM Guillain-Barre syndrome;
 KM chronic inflammatory demyelinating polyneuropathy.
 XX
 XX Homo sapiens.
 XX
 XX 06-NOV-2003; 2003MO-US035268.
 XX 08-NOV-2002; 2002US-0425235P.
 XX (GENE-) GENENTECH INC.
 XX Fong S, Dennis K, Clark H, Chiu H, Schoenfeld J, Williams PM;
 XX Wood WT, Wu TD;
 XX WPI; 2004-420067/39.
 XX P-PSDB; ADO19071.
 XX
 XX Novel PRO polypeptide e.g., PRO69614, PRO71106, or PRO86388 useful for
 PT treating an immune related disorder such as systemic lupus erythematosus,
 PT rheumatoid arthritis, osteoarthritis, juvenile chronic arthritis or
 PT spondyloarthritis.
 PS Claim 1; SEQ ID NO 7; 1731pp; English.
 XX
 XX The invention relates to human PRO polypeptides and the polynucleotides
 CC encoding them. The polypeptides and polynucleotides are useful for
 CC treating and diagnosing immune related disorders in mammals. The immune
 CC related disorders include systemic lupus erythematosus, rheumatoid
 CC arthritis, osteoarthritis, juvenile chronic arthritis, systemic
 CC sclerosis, Sjogren's syndrome, vasculitis, sarcoidosis, autoimmune
 CC haemolytic anaemia, autoimmune thrombocytopenia, thyroiditis, diabetes
 CC mellitus, immune-mediated renal disease, demyelinating polyneuropathy,
 CC central or peripheral nervous system, demyelinating polyneuropathy,
 CC Guillain-Barre syndrome and chronic inflammatory demyelinating
 CC polyneuropathy. This sequence represents a human PRO polynucleotide of
 CC the invention.
 XX
 XX Sequence 2417 BP; 788 A; 503 C; 521 G; 605 T; 0 U; 0 Other;
 SQ
 Query Match 79.0%; Score 15.8; DB 12; Length 2417;
 Best Local Similarity 89.5%; Pred. No. 4.9e+02;
 Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

1 AGTCGCAACAGTACATG 19
 |||||
 DB 2101 AGTCGCAACAGTACATG 2083

RESULT 22
 ADO19457/c
 ID ADO19457 standard; cDNA; 2417 BP.


```

XX AC ADO19457;
XX XX
XX DT 12-AUG-2004 (first entry)
XX DE
XX XX Human PRO polynucleotide #194.
XX XX
XX KW Human; PRO; gene; ss; immune related disorder;
XX KW systemic lupus erythematosus; rheumatoid arthritis; osteoarthritis;
XX KW juvenile chronic arthritis; systemic sclerosis; Sjogren's syndrome;
XX KW vasculitis; sarcoidosis; autoimmune haemolytic anaemia;
XX KW autoimmune thrombocytopenia; thyroiditis; diabetes mellitus;
XX KW renal disease; demyelinating disease; central nervous system;
XX KW peripheral nervous system; demyelinating polyneuropathy;
XX KW Guillain-Barre syndrome;
XX KW chronic inflammatory demyelinating polyneuropathy.
XX OS
XX XX Homo sapiens.
XX XX
XX XX WO2004043361-A2.
XX XX
XX XX 27-MAY-2004.
XX XX
XX XX 06-NOV-2003; 2003WO-US035268.
XX XX
XX XX 08-NOV-2002; 2002US-0425235P.
XX XX
XX XX (GENTH ) GENTECH INC.
XX XX
XX XX Fong S, Dennis K, Clark H, Chiu H, Schoenfeld J, Williams PM,
XX XX Wood W, Wu TD,
XX XX WPI, 2004-420067/39.
XX XX
XX XX P-PSDB; ADO19458.
XX XX
XX XX Novel PRO polypeptide e.g., PRO69614, PRO71106, or PRO8638 useful for
XX XX treating an immune related disorder such as systemic lupus erythematosus,
XX XX rheumatoid arthritis, osteoarthritis, juvenile chronic arthritis or
XX XX spondyloarthritis.
XX XX
XX XX Claim 1; SEQ ID NO 400; 1731bp; English.
XX XX
XX XX The invention relates to human PRO polypeptides and the polynucleotides
XX XX encoding them. The polypeptides and polynucleotides are useful for
XX XX treating and diagnosing immune related disorders in mammals. The immune
XX XX related disorders include systemic lupus erythematosus, rheumatoid
XX XX arthritis, osteoarthritis, juvenile chronic arthritis, systemic
XX XX sclerosis, Sjogren's syndrome, vasculitis, sarcoidosis, autoimmune
XX XX haemolytic anaemia, autoimmune thrombocytopenia, thyroiditis, diabetes
XX XX mellitus, immune-mediated renal disease, demyelinating diseases of the
XX XX central or peripheral nervous system, demyelinating polyneuropathy,
XX XX Guillain-Barre syndrome and chronic inflammatory demyelinating
XX XX polyneuropathy. This sequence represents a human PRO polynucleotide of
XX XX the invention.
XX XX
XX XX Sequence 2417 BP; 788 A; 503 C; 521 G; 605 T; 0 U; 0 Other;
XX XX
XX XX Query Match 79.0%; Score 15.8; DB 12; Length 2417;
XX XX Best Local Similarity 89.5%; Pred. NO. 4.9e+02;
XX XX Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
XX XX
XX QY 1 AGTCGCAACAGTACATG 19
XX DB 2101 AGCTGCAACAGTACATG 2083
XX
XX
XX RESULT 23
XX ID ADF82712/C
XX AD ADF82712 standard; DNA; 2716 BP.
XX AC
XX AD ADF82712;
XX XX
XX XX 26-FEB-2004 (first entry)
XX DT

```

```

XX DE Leukaemia-related DNA sequence #3268.
XX XX
XX KW Cytostatic; Gene therapy; leukaemia; ss.
XX XX
XX OS
XX XX Unidentified.
XX XX
XX XX WO2003039443-A2.
XX XX
XX XX 15-MAY-2003.
XX XX
XX XX 04-NOV-2002; 2002WO-EP012303.
XX XX
XX XX 05-NOV-2001; 2001EP-00126244.
XX XX
XX XX 30-APR-2002; 2002EP-00009758.
XX XX
XX XX (DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.
XX XX (UJLU-) UNIV LUDWIG MAXIMILIANS.
XX XX (HAFE/) HAFERLACH T.
XX XX (SCHO/) SCHOCH C.
XX XX (KERN/) KERN W.
XX XX
XX XX Haferlach T, Schoch C, Kern W, Kohlmann A, Schmittger S, Dugas M,
XX XX Rils R, Broers B, Mergenthaler S,
XX XX WPI, 2003-505037/47.
XX XX
XX XX Determining the subtype of leukemia cells and whether a patient sample
XX XX contains leukemia cells or other cells, useful for treating leukemia,
XX XX comprises determining the expression profile of a group of markers in a
XX XX patient sample.
XX XX
XX XX Disclosure; SEQ ID NO 3268; 2938bp; English.
XX XX
XX XX The present invention relates to a method (M1) for determining the
XX XX CC subtype of leukemia cells and whether a patient sample contains
XX XX CC leukemia cells. The method comprises determining the expression profile
XX XX CC of a group of markers in a patient sample. The method is useful for
XX XX CC determining the presence of leukemia cells, its types or subtypes, and
XX XX CC for the preparation of a medicament for treating leukaemia.
XX XX
XX XX Sequence 2716 BP; 767 A; 463 C; 485 G; 609 T; 0 U; 392 Other;
XX XX
XX XX Query Match 79.0%; Score 15.8; DB 10; Length 2716;
XX XX Best Local Similarity 89.5%; Pred. NO. 4.9e+02;
XX XX Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
XX XX
XX QY 1 AGTCGCAACAGTACATG 19
XX DB 2285 AGCTGCAACAGTACATG 2267
XX
XX
XX RESULT 24
XX ID ABL04367
XX AD ABL04367 standard; CDNA; 2772 BP.
XX AC
XX AB ABL04367;
XX XX
XX XX 26-MAR-2002 (first entry)
XX XX
XX XX Drosophila melanogaster expressed polynucleotide SEQ ID NO 7583.
XX XX
XX KW Drosophila; developmental biology; cell signalling; insecticide;
XX KW pharmaceutical; gene; ss.
XX XX
XX XX Drosophila melanogaster.
XX XX
XX XX WO200171042-A2.
XX XX
XX XX 27-SEP-2001.
XX XX
XX XX 23-MAR-2001; 2001WO-US009231.
XX XX
XX XX

```

PR 23-MAR-2000; 2000US-0191637P.
PR 11-JUL-2000; 2000US-00614150.
XX
PA (PEKE) PE CORP NY.
XX
PI Venter UC, Adams M, Li FMD, Myers BW,
XX
XX WPI; 2001-656860/75.
DR P-PSDB; ABB60264.
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signaling and cell-cell
PT interactions.
XX
PS Claim 1; SEQ ID NO 7583; 21pp + Sequence Listing; English.
XX
CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signaling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (AB116176-AB130511), expressed DNA
CC sequences (AB101840-AB116175) and the encoded proteins (ABB57737-
CC ABB72072). The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 2772 BP; 761 A; 684 C; 689 G; 638 T; 0 U; 0 Other;
XX
Query Match 79.0%; Score 15.8; DB 4; Length 2772;
Best Local Similarity 89.5%; Pred. No. 5e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
XX
QY 1 AGTCGCAACAGTACAAATG 19
Db 493 AATGCAACAGTACAAATG 511
XX
RESULT 25
ADB58306/c
ID ADB58306 standard; DNA; 4183 BP.
XX
AC ADB58306;
XX
DT 04-DEC-2003 (first entry)
XX
DE Toxicity-related gene, SEQ ID 3332.
XX
KW Toxic; toxin; gene expression profile; hepatotoxicity; liver;
KW drug screening; toxicity assay; ds.
XX
OS Unidentified.
XX
PN WO2003064624-A2.
XX
PD 07-AUG-2003.
XX
PF 31-JAN-2003; 2003WO-US003194.
XX
XX 31-JAN-2002; 2002US-00060087.
PR 15-MAR-2002; 2002US-0364045P.
PR 15-MAR-2002; 2002US-0364055P.
PR 30-DEC-2002; 2002US-0436643P.
XX
PA (GENE-) GENE LOGIC INC.
XX
PI Mendrick D, Porter M, Johnson K, Higgs B, Castle A, Elashoff M;
XX
XX WPI; 2003-689530/65.
DR
PT Predicting a toxic effect of a compound, useful in identifying toxicity
PT markers in liver tissues or cells for drug screening and toxicity assays,
PT comprises preparing gene expression profile of tissue or cells exposed to

PT the compound.
XX
PS Claim 1; SEQ ID NO 3332; 1156pp; English.
XX
CC The present invention relates to a method for predicting a toxic effect
CC of a compound. The method comprises preparing a gene expression profile
CC of a tissue or cell sample exposed to the compound, and comparing the
CC gene expression profile to a database comprising SEQ ID 1-4925, where
CC differential expression of the gene indicates at least one toxic effect.
CC The method is useful for predicting at least one toxic effect of a
CC compound, predicting hepatotoxicity or the progression of a toxic effect
CC of a compound, identifying an agent that modulates the onset or
CC progression of a toxic response, predicting the cellular pathways that a
CC compound modulates in a cell, and identifying an agent that modulates at
CC least one activity of a protein. The method and compositions of the
CC present invention using a database of genes having liver toxin-induced
CC differential expression, are useful in identifying toxicity markers in
CC liver tissues or cells for drug screening and toxicity assays. Note: The
CC sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 4183 BP; 1112 A; 1050 C; 926 G; 1095 T; 0 U; 0 Other;
XX
Query Match 79.0%; Score 15.8; DB 10; Length 4183;
Best Local Similarity 89.5%; Pred. No. 5.2e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
XX
QY 2 GTTCGCAACAGTACAAATG 20
Db 3404 GTTCGCAACAGTACAAAGG 3386
XX
RESULT 26
ADB69202
ID ADB69202 standard; DNA; 5238 BP.
XX
AC ADB69202;
XX
DT 04-DEC-2003 (first entry)
XX
DE C. neoformans genomic DNA sequence SEQ ID NO:329.
XX
KW ds; gene; fungicide; gene therapy; infection.
XX
OS Cryptococcus neoformans.
XX
PN WO2003052076-A2.
XX
PD 26-JUN-2003.
XX
PF 17-DEC-2002; 2002WO-US040225.
XX
PR 17-DEC-2001; 2001US-0341261P.
XX
PA (ELIT-) ELITRA PHARM INC.
XX
PI Zamudio C, Eroshkin AM;
XX
XX WPI; 2003-533017/50.
DR P-PSDB; ADB70285.
XX
PT New nucleic acid, useful for preparing a composition for treating an
PT infection caused by Cryptococcus neoformans.
XX
PS Claim 3; SEQ ID NO 329; 136pp; English.
XX
CC The invention relates to a novel purified or isolated Cryptococcus
CC neoformans nucleic acid molecule comprising a sequence encoding a
CC polypeptide comprising a sequence not given in the specification. A
CC polynucleotide of the invention has fungicide activity, and may have a
CC use in gene therapy. The nucleic acid is useful for preparing a
CC composition for treating an infection caused by Cryptococcus neoformans.

CC The present sequence represents a C. neoformans sequence of the
CC invention. Note: the sequence data for this patent is not represented in
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pat_sequences.

XX SQ Sequence 5238 BP; 1147 A; 1479 C; 1211 G; 1400 T; 0 U; 1 Other;

Query Match 79.0%; Score 15.8; DB 10; Length 5238;

Best Local Similarity 89.5%; Pred. No. 5.3e+02;

Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AGTCGCAACAGTACATG 19
|||||
DB 51 AGTCGCAACAGCAGCATG 69

RESULT 27

AAVS2303/c
ID AAVS2303 standard; DNA; 8876 BP.

XX AAVS2303;

XX 23-OCT-1998 (first entry)

DE Streptococcus pneumoniae genome fragment SEQ ID NO:170.

XX Streptococcus pneumoniae; S. pneumoniae; genome; diagnosis; assay;

KW computer readable medium; vaccine; pharmaceutical composition; de.

XX Streptococcus pneumoniae.

OS WO9818931-A2.

XX 07-MAY-1998.

XX 30-OCT-1997; 97WO-US019588.

XX 31-OCT-1996; 96US-0029960P.

PA (HUMA-) HUMAN GENOME SCI INC.

XX Kunsch CA, Choi GH, Dillon PJ, Rosen CA, Barash SC, Fannon M;

PI Dougherty BA;

DR WI; 1998-272225/24.

XX Computer-readable medium with recorded Streptococcus pneumoniae

PT polynucleotide sequences - useful in diagnostic kits and assays, and

PT pharmaceutical compositions and vaccines for Streptococcus pneumoniae.

PS Claim 1; Page 1080-1085; 1409PP; English.

XX The present invention describes a computer readable medium which has the
CC nucleotide sequences SEQ ID NO:1 to 391 (AAVS2134 to AAVS2524) recorded
CC on it, or a representative fragment or a sequence at least 95% identical
CC to SEQ ID NO:1 to 391. The nucleotide sequences depicted in SEQ ID NO:1
CC to 391 (AAVS2134 to AAVS2524) are genomic fragments from Streptococcus
CC pneumoniae. The present invention also describes an isolated nucleic acid
CC molecule encoding a homologue of any of the fragments of the S. pneumoniae
CC genome (SEQ ID NO:1 to 391) where the nucleic acid molecule is produced
CC by a process comprising: (a) screening a genomic DNA library using as a
CC probe a target sequence defined by any of the sequences in SEQ ID NO:1 to
CC 391, identifying members of the library which contain sequences that
CC hybridize to the target sequence and isolating the nucleic acid molecules
CC from the members; or (b) isolating mRNA, DNA or cDNA produced from an
CC organism, amplifying nucleic acid molecules whose nucleotide sequence is
CC homologous to amplification primers derived from the fragment of the S.
CC pneumoniae genome to prime the amplification and isolating the amplified
CC sequences. The computer readable medium can be used in a computer-based
CC system for identifying fragments of the S. pneumoniae genome of
CC pneumoniae importance, or expression modulating fragments of the S.
CC pneumoniae genome. Products from the present invention can be used in
CC diagnosis kits and assays, and pharmaceutical compositions and vaccines

CC for S. pneumoniae

XX SQ Sequence 8876 BP; 2663 A; 1840 C; 1505 G; 2867 T; 0 U; 1 Other;

Query Match 79.0%; Score 15.8; DB 2; Length 8876;

Best Local Similarity 89.5%; Pred. No. 5.7e+02;

Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GTCGCAACAGTACATG 20
|||||
DB 4731 GTCGCAACAGCAGCATG 4713

RESULT 28

ABS56454_07
Continuation (8 of 22) of ABS56454 from base 700001 (Streptococcus pneumoniae type 4 str:

WP Sequence split into 22 fragments LOCUS ABS56454 Accession ABS56454

WP Fragment Name Begin End

WP ABS56454_00 1 11000

WP ABS56454_01 100001 210000

WP ABS56454_02 200001 310000

WP ABS56454_03 300001 410000

WP ABS56454_04 400001 510000

WP ABS56454_05 500001 610000

WP ABS56454_06 600001 710000

WP ABS56454_07 700001 810000

WP ABS56454_08 800001 910000

WP ABS56454_09 900001 1010000

WP ABS56454_10 1000001 1110000

WP ABS56454_11 1100001 1210000

WP ABS56454_12 1200001 1310000

WP ABS56454_13 1300001 1410000

WP ABS56454_14 1400001 1510000

WP ABS56454_15 1500001 1610000

WP ABS56454_16 1600001 1710000

WP ABS56454_17 1700001 1810000

WP ABS56454_18 1800001 1910000

WP ABS56454_19 1900001 2010000

WP ABS56454_20 2000001 2110000

WP ABS56454_21 2100001 2162598

Query Match 79.0%; Score 15.8; DB 10; Length 110000;

Best Local Similarity 89.5%; Pred. No. 7.5e+02;

Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GTCGCAACAGTACATG 20
|||||

DB 52380 GTCGCAACAGCAGCATG 52398

XX RESULT 29

ADQ20017

XX ADQ20017 standard; DNA; 260160 BP.

XX ADQ20017;

XX 26-AUG-2004 (first entry)

DE Human soft tissue sarcoma-upregulated DNA - SEQ ID 2837.

XX soft tissue sarcoma; cytostatic; gene therapy; vaccine; screening; human;

KW ds.

XX Homo sapiens.

OS WO2004048938-A2.

XX 10-JUN-2004.

XX 26-NOV-2003; 2003WO-US038193.

XX 26-NOV-2002; 2002US-0429739P.

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

PA (PROT-) PROTEIN DESIGN LABS INC.
XX
PI Aziz N, Ginsburg WM, Zlotnik A;
XX
DR WPI; 2004-441208/41.
XX
PT Early detection of soft tissue sarcoma comprises determining expression
PT of a gene in a first soft tissue sample and a normal soft tissue sample
PT and comparing the gene expression, also useful in treating soft tissue
PT sarcoma.
XX
PS Example 2; SEQ ID NO 2837; 210pp; English.
XX
CC The invention relates to a novel method for detecting soft tissue sarcoma
CC which comprises obtaining a first soft tissue sample from an individual
CC and a normal soft tissue sample from the same or different individual,
CC determining the expression of a gene in both samples and comparing the
CC expression of the gene in both soft tissue samples, where a higher level
CC of protein expression in the first soft tissue sample indicates the
CC presence of soft tissue sarcoma. The method of the invention has
CC cyostatic applications and may be useful for detecting soft tissue
CC sarcoma, possibly via gene therapy or vaccine production. The nucleic
CC acid sequences may be useful in diagnostic and screening applications.
CC The current sequence is that of a human soft tissue sarcoma-upregulated
CC DNA of the invention. The current sequence is not shown within the
CC specification per se but was submitted in CD format by the inventor.
XX
SQ Sequence 260160 BP; 68455 A; 56280 C; 57380 G; 78045 T; 0 U; 0 Other;
XX
QY
Query Match 79.0%; Score 15.8; DB 12; Length 260160;
Best Local Similarity 89.5%; Pred. NO. 8.2e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
1 AGTGCACAAACAGTACATG 19
152805 AATGCGAAACAGTACATG 152823
XX
RESULT 30
ADB82581/C
ID ADB82581 standard; cDNA; 236 BP.
XX
AC ADB82581;
XX
DT 04-DEC-2003 (first entry)
XX
DE Human cDNA sequence useful for the treatment of cancer (SeqID 893).
XX
KW human; prostate; cancer; cytostatic; gene therapy; vaccine;
KW immune response; gene; ss.
XX
OS Homo sapiens.
OS
PN WO2003050236-A2.
XX
PD 19-JUN-2003.
XX
PF 04-SEP-2002; 2002WO-US028214.
XX
PR 07-DEC-2001; 2001US-00012697.
XX
PA (CHIR) CHIRON CORP.
PA (HYSE-) HYSEQ INC.
XX
PI Escobedo J, Garcia PD, Kassam A, Lamson G, Drmanac R;
PI Cirkvenjakov R, Dickson M, Drmanac S, Labat I, Leshkowitz D, Kita D;
PI Garcia V, Jones LW, Stache-Crain B, Scott EM;
XX
DR WPI; 2003-513972/48.
XX
PT New polynucleotides derived from human prostate, useful for modulating
PT immune response to prevent or treat cancer.
XX

PS Claim 1; SEQ ID NO 893; 188pp; English.
XX
CC This invention relates to novel isolated polynucleotides of human origin,
CC particularly isolated from the human prostate. Specifically, it refers to
CC the diagnostics and therapeutics comprising these novel human
CC polynucleotides, and includes the derived probes, antisense
CC oligonucleotides and antibodies thereof. The identification of these
CC human prostate genes that can inhibit tumour growth is useful for
CC understanding the progression and nature of complex diseases such as
CC cancer, and hence they are important in the drug discovery process. The
CC present invention describes these polynucleotides and encoded
CC polypeptides as exhibiting cytostatic activity, and through gene therapy
CC and/or vaccines they can be used to modulate the immune response for the
CC prevention or treatment of cancers, particularly of the prostate, but
CC also for breast, lung and colon cancer. This polynucleotide sequence is a
CC human cDNA sequence useful for the treatment of cancer, used in an
CC exemplification of the invention. NOTE: These sequences are not given in
CC the specification but are provided on the WIPO website.
XX
SQ Sequence 236 BP; 66 A; 41 C; 54 G; 49 T; 0 U; 26 Other;
XX
QY
Query Match 77.0%; Score 15.4; DB 9; Length 236;
Best Local Similarity 94.1%; Pred. NO. 6e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
1 AGTGCACAAACAGTACCA 17
202 AGTGCACAAACAGTACCA 186
XX
DB
RESULT 31
ABQ89640/C
ID ABQ89640 standard; cDNA; 306 BP.
XX
AC ABQ89640;
XX
DT 27-SEP-2002 (first entry)
XX
DE Human prostate expressed polynucleotide SEQ ID NO 896.
XX
KW Human; prostate; cytostatic; tumour; cancer; vaccine; gene therapy; gene;
KW ss.
XX
OS Homo sapiens.
OS
PN WO200255700-A2.
XX
PD 18-JUL-2002.
XX
PF 07-DEC-2001; 2001WO-US047349.
XX
PR 07-DEC-2000; 2000US-0254648P.
PR 13-MAR-2001; 2001US-0275688P.
XX
PA (CHIR) CHIRON CORP.
PA (HYSE-) HYSEQ INC.
XX
PI Escobedo J, Garcia PD, Kassam A, Lamson G, Drmanac R;
PI Cirkvenjakov R, Dickson M, Drmanac S, Labat I, Leshkowitz D, Kita D;
PI Garcia V, Jones WL, Stache-Crain B, Scott EM;
XX
DR WPI; 2002-557824/59.
XX
PT New genes and gene products isolated from human prostate, useful for
PT treating or diagnosing tumor or cancer (e.g. prostate cancer or breast
PT cancer), or as vaccines for treating or preventing these diseases.
XX
PS Claim 1; SEQ ID NO 896; 186pp + Sequence Listing; English.
XX
CC The invention relates to an isolated polynucleotide comprising any of
CC 1477 sequences or its fragment, degenerate variant, antisense or
CC complement. The polynucleotides and gene products are useful for treating
CC or diagnosing tumour or cancer (e.g. prostate cancer, breast cancer, lung

CC cancer or medullary carcinoma) in a subject (e.g. cattle, dogs, cats,
CC rabbits, horse or human). The polynucleotides and polypeptides are also
CC useful as vaccines for treating or preventing these diseases. The
CC polynucleotides are useful for gene therapy. The present sequence is that
CC of one of a group of polynucleotides (ABQ8745-ABQ9015) disclosed
CC electronically as sequences of the invention. However only 1271
CC polynucleotide sequences are given, whereas 1477 polynucleotides and 91
CC proteins are claimed. Note: The sequence data for this patent did not
CC form part of the printed specification, but was obtained in electronic
CC format directly from WIPO at ftp.wipo.int/pub/published_pcc_sequence

XX
SQ Sequence 306 BP; 83 A; 55 C; 76 G; 61 T; 0 U; 31 Other;

Query Match 77.0%; Score 15.4; DB 6; Length 306;
Best Local Similarity 94.1%; Pred. No. 6.1e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AGTGCACAAACAGTACAA 17
DB 272 AGTGCACAAACAGTACCA 256

RESULT 32
ABV77453/C
ID ABV77453 standard; cDNA; 306 BP.
XX
AC ABV77453;
XX
DT 31-JAN-2003 (first entry)
XX
DE P. monodon full-length putative reproductive inhibiting hormone cDNA.
XX
KW Reproductive inhibiting hormone; spawning; penaeid; prawn; contraceptive;
KW antifertility; gene therapy; SGPVI; RH; PmSGPVI; gene; ss.
XX
OS Penaeus monodon.
XX
FH Key Location/Qualifiers
FT 1..306
FT CDS /*tag= b
FT /product= "Reproductive inhibiting hormone"
FT /note= "No stop codon given"
FT sig_peptide 1..81
FT /*tag= c
FT mat_peptide 82..306
FT /*tag= d

WO200283717-A1.
24-OCT-2002.
XX
PD 11-APR-2002; 2002MO-AU000466.
XX
PF 11-APR-2001; 2001AU-00004368.
XX
PR 16-MAY-2001; 2001AU-00005049.
XX
PT (AUMA-) AUSTRALIAN INST MARINE SCI.
XX
PI Davey ML, Hall MR, Swan JA, Wilson KJ;
XX
DR WPI; 2003-092997/08.
XX
DR P-PSDB; ABP57824, ABP57825.
XX
PT New peptides from Penaeus monodon, useful for inhibiting spawning in
PT penaeid prawns, or for screening inhibitors or antagonists of the
PT reproductive inhibiting hormone, which are useful for inducing spawning
PT in Crustacea.
XX
XX
PS Claim 9; Page 6; 70pp; English.
XX
CC The invention relates to a novel isolated polypeptide from Penaeus
CC monodon, an active fragment of the polypeptide, or a peptide with
CC substantial sequence identity to the polypeptide, which serves to inhibit

CC spawning in penaeid prawns. The protein of the invention has
CC contraceptive and antifertility activity. The invention may have a use
CC in gene therapy. The protein is useful for inhibiting spawning in penaeid
CC prawns. This peptide is also useful for screening inhibitors of RH. The
CC antibody or RNAi is useful as an antagonist of a reproductive inhibiting
CC hormone, and is useful for inducing spawning in Crustacea. The present
CC sequence encodes the P. monodon putative reproductive inhibiting hormone
CC (RH), also referred to as PmSGPVI or SGPVI

XX
SQ Sequence 306 BP; 78 A; 60 C; 85 G; 83 T; 0 U; 0 Other;

Query Match 77.0%; Score 15.4; DB 8; Length 306;
Best Local Similarity 94.1%; Pred. No. 6.1e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GTCCGAAACAGTACAAAT 18
DB 56 GTCCGAAATAGTACAAAT 40

RESULT 33
ACH49767/C
ID ACH49767 standard; cDNA; 436 BP.
XX
AC ACH49767;
XX
DT 13-OCT-2003 (first entry)
XX
DE Human leukocyte cDNA #1361.
XX
KW Human; ss; sequencing by hybridisation; SHH; expressed sequence tag; EST;
KW genome mapping; biodiversity; genetic disorder.
XX
OS Homo sapiens.
XX
PN US2003073623-A1.
XX
PD 17-APR-2003.
XX
PF 30-JUL-2001; 2001US-00918995.
XX
PR 30-JUL-2001; 2001US-00918995.
XX
PA (DRMA/) DRMANAC R T.
PA (STRAC/) STRACH-CRAIN B.
PA (DICK/) DICKSON M C.
PA (JONE/) JONES L W.
XX
PI Drmanac RT, Labat I, Stache-Crain B, Dickson MC, Jones LW;
XX
DR WPI; 2003-615964/58.
XX
PT New polynucleotide sequences obtained from various cDNA libraries, useful
PT as hybridization probes, as oligomers for PCR, for chromosome and gene
PT mapping, in the recombinant production of protein, or in generating
PT antisense DNA or RNA.
XX
XX
PS Claim 1; SEQ ID NO 36979; 44pp; English.
XX
CC The invention relates to an isolated polynucleotide comprising any one of
CC 38043 cDNA sequences, appearing as ACH12789-ACH50831, whose sequence was
CC determined by the technique of SBH (sequencing by hybridisation). Also
CC included is a purified polypeptide comprising a sequence corresponding to
CC a reading frame of the novel polynucleotide. The nucleic acid sequences
CC are useful in diagnostics as expressed sequence tags (EST) for
CC identifying expressed genes or for physical mapping of the human genome,
CC in forensics, in assessing biodiversity, or in identifying mutations
CC responsible for genetic disorders and other traits. The nucleotide
CC sequences are also useful as hybridisation probes, as oligomers for PCR,
CC for chromosome and gene mapping, in the recombinant production of
CC protein, or in generating antisense DNA or RNA. The purified polypeptide
CC is useful for generating antibodies specific for it. The present sequence

CC is one of the 38043 isolated cDNA/EST sequences. Note: The sequence data
CC for this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from USPTO at
CC seqdata.uspto.gov/sequence.html?docid=20030073623

XX Sequence 436 BP; 140 A; 79 C; 86 G; 131 T; 0 U; 0 Other;

QY Query Match 77.0%; Score 15.4; DB 9; Length 436;
Best Local Similarity 94.1%; Pred. No. 6.4e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

XX 4 CGCAACAGTACATGG 20
DB 23 CGCAACAGTACATAG 7

RESULT 34
ABZ73091
ID ABZ73091 standard; cDNA; 499 BP.

XX ABZ73091;

XX 10-APR-2003 (first entry)

XX Rice leaf EST, SEQ ID NO:29.

XX Rice; leaf; EST; expressed sequence tag; plant; biochip; DNA array;
XX cloning; detection; heterosis; hybrid vigour; transgene detection;
XX herbicide screening; pesticide screening; disease diagnosis;
XX medical research; agriculture; ss.

XX Oryza sativa.

XX CN1364936-A.

XX 21-AUG-2002.

XX 31-OCT-2001; 2001CN-00137672.

XX 31-OCT-2001; 2001CN-00137672.

XX (UYZH-) UNIV ZHEJIANG.

XX Li D, Dong H;

XX WPI; 2003-000550/01.

XX New rice leaf expression sequence labels and constituted biochip.

XX Claim 1; Page 20 (disclosure); 29pp; Chinese.

XX The invention relates to 50 novel ESTs (expressed sequence tags; ABZ73063
XX -ABZ73113) obtained from a rice leaf cDNA library. The invention also
XX relates to a biochip comprising these ESTs. The biochip of the invention
XX can be used in a variety of agricultural applications. It can be used in
XX the cloning of genes which confer useful characteristics in crop species,
XX to predict hybrid vigour (heterosis) at an early stage, in the detection
XX of transgenic agricultural produce, in screening for novel herbicides and
XX pesticides and in disease diagnosis. The biochip may also be used in
XX medical research. The present sequence represents a rice leaf EST of the
XX invention

XX Sequence 499 BP; 116 A; 128 C; 132 G; 123 T; 0 U; 0 Other;

QY Query Match 77.0%; Score 15.4; DB 8; Length 499;
Best Local Similarity 94.1%; Pred. No. 6.5e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 CGCAACAGTACATGG 20
DB 319 CGCAACAGTACATGG 335

RESULT 35
ABA60404
ID ABA60404 standard; DNA; 542 BP.

XX ABA60404;

XX 01-FEB-2002 (first entry)

XX Human foetal liver single exon nucleic acid probe #8709.

XX Human; foetal liver; gene expression; single exon nucleic acid probe; ss.

XX Homo sapiens.

XX WO200157277-A2.

XX 09-AUG-2001.

XX 30-JAN-2001; 2001WO-US000669.

XX 04-FEB-2000; 2000US-0180312P.

XX 26-MAY-2000; 2000US-0207456P.

XX 30-JUN-2000; 2000US-00608408.

XX 03-AUG-2000; 2000US-00632366.

XX 21-SEP-2000; 2000US-0234687P.

XX 27-SEP-2000; 2000US-0236359P.

XX 04-OCT-2000; 2000GB-00024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2001-483447/52.

XX Human genome-derived single exon nucleic acid probes useful for analyzing

XX gene expression in human fetal liver.

XX Claim 1; SEQ ID NO 8709; 639pp + Sequence Listing; English.

XX The invention relates to a single exon nucleic acid probe for measuring

XX human gene expression in a sample derived from human foetal liver. The

XX single exon nucleic acid probes may be used for predicting, measuring and

XX displaying gene expression in samples derived from human fetal liver. The

XX present sequence is a single exon nucleic acid probe of the invention.

XX Note: The sequence data for this patent did not form part of the printed

XX specification, but was obtained in electronic format directly from WIPO

XX at ftp.wipo.int/pub/published_pct_sequences

XX Sequence 542 BP; 154 A; 147 C; 122 G; 119 T; 0 U; 0 Other;

QY Query Match 77.0%; Score 15.4; DB 4; Length 542;
Best Local Similarity 94.1%; Pred. No. 6.6e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 CGCAACAGTACATGG 20
DB 58 CGCAACAGTACATAG 74

RESULT 36
AA140288
ID AA140288 standard; DNA; 542 BP.
XX AA140288;
XX 17-OCT-2001 (first entry)
XX Probe #8974 used to measure gene expression in human placenta sample.
XX Probe; microarray; human; placenta; antenatal diagnosis;
XX genetic disorder; ss.
XX Homo sapiens.

```
XX XX WO200157272-A2.
XX XX
XX PD 09-AUG-2001.
XX PF
XX PF 30-JAN-2001; 2001WO-US000663.
XX PR 04-FEB-2000; 2000US-0180312P.
XX PR 26-MAY-2000; 2000US-0207456P.
XX PR 30-JUN-2000; 2000US-00608408.
XX PR 03-AUG-2000; 2000US-00632366.
XX PR 21-SEP-2000; 2000US-0234687P.
XX PR 27-SEP-2000; 2000US-0236359P.
XX PR 04-OCT-2000; 2000GB-00024263.
XX PA
XX PA (MOLE-) MOLECULAR DYNAMICS INC.
XX PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX DR WPI; 2001-488897/53.
XX PT Human genome-derived single exon nucleic acid probes useful for analyzing
XX PT gene expression in human placenta.
XX PS Claim 25; SEQ ID NO 8974; 654pp; English.
XX CC The present invention relates to single exon nucleic acid probes (SNP).
XX CC The present sequence is one such probe. The probes are useful for
XX CC producing a microarray for predicting, measuring and displaying gene
XX CC expression in samples derived from human placenta. The probes are useful
XX CC for antenatal diagnosis of human genetic disorders
XX SQ Sequence 542 BP; 154 A; 147 C; 122 G; 119 T; 0 U; 0 Other;
XX
XX Query Match 77.0%; Score 15.4; DB 4; Length 542;
XX Best Local Similarity 94.1%; Pred. No. 6.6e+02;
XX Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 4 CGCAACAGTACATG 20
DB 58 CGCAACAGTACATAG 74
XX
XX RESULT 37
XX ABA28630
XX ID ABA28630 standard; DNA; 542 BP.
XX AC ABA28630;
XX XX
XX DT 23-JAN-2002 (first entry)
XX DE Probe #7096 for gene expression analysis in human heart cell sample.
XX KW Human; gene expression; heart; microarray; vascular system; probe;
XX KW cardiovascular disease; hypertension; cardiac arrhythmia;
XX KW congenital heart disease; ss.
XX OS Homo sapiens.
XX PN WO200157274-A2.
XX PD 09-AUG-2001.
XX PF 30-JAN-2001; 2001WO-US000666.
XX PR 04-FEB-2000; 2000US-0180312P.
XX PR 26-MAY-2000; 2000US-0207456P.
XX PR 30-JUN-2000; 2000US-00608408.
XX PR 03-AUG-2000; 2000US-00632366.
XX PR 21-SEP-2000; 2000US-0234687P.
XX PR 27-SEP-2000; 2000US-0236359P.
XX PR 04-OCT-2000; 2000GB-00024263.
XX
```

```
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX DR WPI; 2001-488899/53.
XX PT Single exon nucleic acid probes for analyzing gene expression in human
XX PT hearts.
XX PS Claim 1; SEQ ID NO 7096; 530pp; English.
XX CC The present invention relates to single exon nucleic acid probes for
XX CC measuring human gene expression in a sample derived from human heart. The
XX CC present sequence is one such probe. The probes may be used for
XX CC predicting, measuring and displaying gene expression in samples derived
XX CC from the human heart via microarrays. By measuring gene expression, the
XX CC probes are useful for predicting, diagnosing, grading, staging,
XX CC monitoring and prognosing diseases of the human heart and vascular system
XX CC e.g. cardiovascular disease, hypertension, cardiac arrhythmias and
XX CC congenital heart disease. Note: The sequence data for this patent did not
XX CC form part of the printed specification, but was obtained in electronic
XX CC format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX SQ Sequence 542 BP; 154 A; 147 C; 122 G; 119 T; 0 U; 0 Other;
XX
XX Query Match 77.0%; Score 15.4; DB 4; Length 542;
XX Best Local Similarity 94.1%; Pred. No. 6.6e+02;
XX Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 4 CGCAACAGTACATG 20
DB 58 CGCAACAGTACATAG 74
XX
XX RESULT 38
XX AAK34570
XX ID AAK34570 standard; DNA; 542 BP.
XX AC AAK34570;
XX XX
XX DT 06-NOV-2001 (first entry)
XX DE Human bone marrow expressed single exon probe SEQ ID NO: 9127.
XX KW Human; bone marrow expressed exon; gene expression analysis; probe;
XX KW microarray; cancer; leukaemia; lymphoma; myeloma; ss.
XX OS Homo sapiens.
XX PN WO200157276-A2.
XX PD 09-AUG-2001.
XX PF 30-JAN-2001; 2001WO-US000668.
XX PR 04-FEB-2000; 2000US-0180312P.
XX PR 26-MAY-2000; 2000US-0207456P.
XX PR 30-JUN-2000; 2000US-00608408.
XX PR 03-AUG-2000; 2000US-00632366.
XX PR 21-SEP-2000; 2000US-0234687P.
XX PR 27-SEP-2000; 2000US-0236359P.
XX PR 04-OCT-2000; 2000GB-00024263.
XX PA
XX PA (MOLE-) MOLECULAR DYNAMICS INC.
XX PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX DR WPI; 2001-488900/53.
XX PT Human genome-derived single exon nucleic acid probes useful for analyzing
XX PT gene expression in human bone marrow.
XX PS Example 4; SEQ ID NO 9127; 658pp + Sequence listing; English.
```


XX CC The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
CC bone marrow. They can be used to measure gene expression in bone marrow
CC samples, which may enable the improved diagnosis and treatment of cancers
CC such as lymphoma, leukaemia and myeloma. The present sequence is one of
CC the probes of the invention
XX
SQ Sequence 542 BP; 154 A; 147 C; 122 G; 119 T; 0 U; 0 Other;
XX
Query Match 77.0%; Score 15.4; DB 4; Length 542;
Best Local Similarity 94.1%; Pred. No. 6.6e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 4 CGCAACAGTACATGG 20
Db 58 CGCAACAGTACATAG 74
|||||
RESULT 39
AAK08682
ID AAK08682 standard; DNA; 542 BP.
XX
AC AAK08682;
XX
DT 05-NOV-2001 (first entry)
XX
DE Human brain expressed single exon probe SEQ ID NO: 8673.
XX
XX Human; brain expressed exon; gene expression analysis; probe; microarray;
XX Alzheimer's disease; multiple sclerosis; schizophrenia; epilepsy; cancer;
XX ss.
XX Homo sapiens.
XX
XX WO200157275-A2.
XX
XX 09-AUG-2001.
XX
XX 30-JAN-2001; 2001WO-US000667.
XX
XX 04-FEB-2000; 2000US-0180312P.
XX 26-MAY-2000; 2000US-0207456P.
XX 30-JUN-2000; 2000US-00608408.
XX 03-AUG-2000; 2000US-00632366.
XX 21-SEP-2000; 2000US-0234687P.
XX 27-SEP-2000; 2000US-0236359P.
XX 04-OCT-2000; 2000GB-00024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX
XX WPI; 2001-483446/52.
XX
XX Single exon nucleic acid probes for analyzing gene expression in human
XX brains.
XX
XX Example 4; SEQ ID NO 8673; 650bp + Sequence Listing; English.
XX
XX The present invention provides a number of single exon nucleic acid
XX probes which are derived from genomic sequences expressed in the human
XX brain. They can be used to measure gene expression in brain cell samples,
XX which may enable the diagnosis and improved treatment of nervous system
XX diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
XX epilepsy and cancers. The present sequence is one of the probes of the
XX invention
XX
SQ Sequence 542 BP; 154 A; 147 C; 122 G; 119 T; 0 U; 0 Other;
XX
Query Match 77.0%; Score 15.4; DB 4; Length 542;
Best Local Similarity 94.1%; Pred. No. 6.6e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 CGCAACAGTACATGG 20
Db 58 CGCAACAGTACATAG 74
|||||
RESULT 40
ABS34344
ID ABS34344 standard; DNA; 542 BP.
XX
AC ABS34344;
XX
DT 25-FEB-2003 (first entry)
XX
XX Human liver single exon probe, SEQ ID No 9334.
XX
XX Human; single exon nucleic acid probe; liver; cirrhosis;
XX hyperlipoproteinaemia; hyperlipidaemia; hypercholesterolaemia;
XX coronary heart disease; ss.
XX
XX Homo sapiens.
XX
XX WO200157273-A2.
XX
XX 09-AUG-2001.
XX
XX 30-JAN-2001; 2001WO-US000664.
XX
XX 04-FEB-2000; 2000US-0180312P.
XX 26-MAY-2000; 2000US-0207456P.
XX 30-JUN-2000; 2000US-00608408.
XX 03-AUG-2000; 2000US-00632366.
XX 21-SEP-2000; 2000US-0234687P.
XX 27-SEP-2000; 2000US-0236359P.
XX 04-OCT-2000; 2000GB-00024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX
XX WPI; 2001-488898/53.
XX
XX Human genome-derived single exon nucleic acid probes useful for analyzing
XX gene expression in human adult liver.
XX
XX Claim 1; SEQ ID NO 9334; 656bp; English.
XX
XX The invention relates to a single exon nucleic acid probe (SENP) (I) for
XX measuring human gene expression in a sample derived from human adult
XX liver, comprising one of 1109 defined nucleotide sequences given in the
XX specification (or complements/ fragments). The probe hybridises at high
XX stringency to a nucleic acid molecule expressed in the human adult liver.
XX (I) may be used for predicting, measuring and displaying gene expression
XX in samples derived from human adult liver. The genes identified may be
XX involved in genetic liver diseases such as cirrhosis,
XX hyperlipoproteinaemia, hyperlipidaemia and hypercholesterolaemia which is
XX associated with coronary heart disease. ABS25011-ABS51005 represent human
XX liver single exon nucleic acid probes of the invention. Note: The
XX sequence information for this patent does not appear in the printed
XX specification but was obtained in electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 542 BP; 154 A; 147 C; 122 G; 119 T; 0 U; 0 Other;
XX
Query Match 77.0%; Score 15.4; DB 4; Length 542;
Best Local Similarity 94.1%; Pred. No. 6.6e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 4 CGCAACAGTACATGG 20
Db 58 CGCAACAGTACATAG 74
|||||

RESULT 41
 ABS09204
 ID ABS09204 standard; DNA: 542 BP.
 AC ABS09204;
 XX
 DT 19-AUG-2002 (first entry)
 DE Human genome-derived single exon probe from lung SEQ ID NO 9195.
 XX
 KW Human; ds; single exon probe; asphma; lung cancer; COPD; ILD;
 KW chronic obstructive pulmonary disease; interstitial lung disease;
 KW familial idiopathic pulmonary fibrosis; neurofibromatosis;
 KW tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;
 KW Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;
 KW pulmonary histiocytosis; lymphangioleiomyomatosis; Karsener syndrome;
 KW pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;
 KW primary ciliary dyskinesia; pulmonary hypertension;
 KW hyaline membrane disease.
 KW
 OS Homo sapiens.
 XX
 PN MO200186003-A2.
 XX
 PD 15-NOV-2001.
 XX
 PF 30-JAN-2001; 2001MO-US000665.
 XX
 PR 04-FEB-2000; 2000US-0180312P.
 PR 26-MAY-2000; 2000US-0207456P.
 PR 30-JUN-2000; 2000US-00608408.
 PR 03-AUG-2000; 2000US-00632366.
 PR 21-SEP-2000; 2000US-0234687P.
 PR 27-SEP-2000; 2000US-0236359P.
 PR 04-OCT-2000; 2000GB-00024263.
 XX
 PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX
 PI Penn SG, Hanzel DK, Chen W, Rank DR;
 XX
 DR WPI; 2002-114183/15.
 XX
 PT Spatially-addressable set of single exon nucleic acid probes, used to
 PT measure gene expression in human lung samples.
 XX
 PS Claim 1; SEQ ID NO 9195; 634bp; English.
 XX
 CC The invention relates to a spatially-addressable set of single exon
 CC nucleic acid probes for measuring gene expression in a sample derived
 CC from human lung comprising single exon nucleic acid probes having one of
 CC 12614 nucleic acid sequences mentioned in the specification, or their
 CC complements or the 12387 open reading frames derived from the 12614
 CC probes. Also included are a microarray comprising the novel set of probes
 CC; the novel set of probes which hybridise at high stringency to a nucleic
 CC acid expressed in the human lung; measuring gene expression in a sample
 CC derived from human lung, comprising (a) contacting the array with a
 CC collection of detectably labeled nucleic acids derived from human lung
 CC mRNA, and (b) measuring the label detectably bound to each probe of the
 CC array; identifying exons in a eukaryotic genome, comprising (a)
 CC algorithmically predicting at least one exon from genomic sequences of
 CC the eukaryote; and (b) detecting specific hybridisation of detectably
 CC labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,
 CC having a fragment identical to the predicted exon, the probe is included
 CC in the above mentioned microarray; assigning exons to a single gene,
 CC comprising (a) identifying exons from genomic sequence by the method
 CC above and (b) measuring the expression of each of the exons in several
 CC tissues and/or cell types using hybridisation to a single exon
 CC microarrays having a probe with the exon, where a common pattern of
 CC expression of the exons in the tissues and/or cell types indicates that
 CC the exons should be assigned to a single gene; a peptide comprising one
 CC of 12011 sequences, mentioned in the specification, or encoded by the
 CC probes/open reading frames (ORF). The probes are used for gene expression
 CC analysis, and for identifying exons in a gene, particularly using human

CC lung derived mRNA and for the study of lung diseases such as asthma, lung
 CC cancer, chronic obstructive pulmonary disease (COPD), interstitial lung
 CC disease (ILD), familial idiopathic pulmonary fibrosis, neurofibromatosis,
 CC tuberous sclerosis, Gaucher's disease, Niemann-Pick disease, Hermansky-
 CC Pudlak syndrome, sarcoidosis, pulmonary haemosiderosis, pulmonary
 CC histiocytosis, lymphangioleiomyomatosis, pulmonary alveolar proteinosis,
 CC Karsener syndrome, fibrocystic pulmonary dysplasia, primary ciliary
 CC dyskinesia, pulmonary hypertension and hyaline membrane disease. The
 CC present sequence is a single exon probe of the invention. Note: The
 CC sequence data for this parent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WPI
 CC at ftp.wpi.edu/pub/published_pcr_sequences
 XX
 SQ Sequence 542 BP; 154 A; 147 C; 122 G; 119 T; 0 U; 0 Other;
 XX
 QY
 DB 4 CGCAAAACGATCAATGG 20
 58 CGCAAAACGATCAATAG 74
 XX
 RESULT 42
 ID ABV77452/C
 ABV77452 standard; CDNA: 818 BP.
 XX
 AC ABV77452;
 XX
 DT 31-JAN-2003 (first entry)
 XX
 DE P. monodon full-length putative reproductive inhibiting hormone cDNA.
 XX
 KW Reproductive inhibiting hormone; spawning; penaeid; prawn; contraceptive;
 KW antifertility; gene therapy; SGPV; RH; PMSGPV; gene; ss.
 XX
 OS Penaeus monodon.
 XX
 FH Key Location/Qualifiers
 FT 5'UTR 1..65
 FT CDS /*tag= a
 FT /*tag= b
 FT /*tag= c
 FT sig_peptide 66..146
 FT mat_peptide 147..371
 FT /*tag= d
 FT 3'UTR 375..818
 FT /*tag= e
 XX
 PN WO200283717-A1.
 XX
 PD 24-OCT-2002.
 XX
 PF 11-APR-2002; 2002MO-AU000466.
 XX
 PR 11-APR-2001; 2001AU-00004368.
 PR 16-MAY-2001; 2001AU-00005049.
 XX
 PA (AUMA-) AUSTRALIAN INST MARINE SCI.
 XX
 PI Davey ML, Hall MR, Swan JA, Wilson KJ;
 XX
 DR WPI; 2003-092997/08.
 DR P-PSDB; ABP57824, ABP57825.
 XX
 CC New peptides from Penaeus monodon, useful for inhibiting spawning in
 CC penaeid prawns, or for screening inhibitors or antagonists of the
 CC reproductive inhibiting hormone, which are useful for inducing spawning
 CC in Crustacea.

PS Claim 8, Page 5-6; 70pp; English.
XX
CC The invention relates to a novel isolated polypeptide from Penaeus
CC monodon, an active fragment of the polypeptide, or a peptide with
CC substantial sequence identity to the polypeptide, which serves to inhibit
CC spawning in penaeid prawns. The protein of the invention has
CC contraceptive and antifertility activity. The invention may have a use
CC in gene therapy. The protein is useful for inhibiting spawning in penaeid
CC prawns. This peptide is also useful for screening inhibitors of RH. The
CC antibody or RNAi is useful as an antagonist of a reproductive inhibiting
CC hormone, and is useful for inducing spawning in Crustacea. The present
CC sequence encodes the P. monodon putative reproductive inhibiting hormone
CC (RH), also referred to as PmsGPVI or SGPVI
XX
SQ Sequence 818 BP; 237 A; 163 C; 172 G; 246 T; 0 U; 0 Other;
Query Match 77.0%; Score 15.4; DB 8; Length 818;
Best Local Similarity 94.1%; Pred. No. 6.9e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 2 GTGGCAACAGTACAAT 18
Db 121 GTGGCAAAATAGTACAAT 105
RESULT 43
AAS77840/C
ID AAS77840 standard; cDNA; 854 BP.
XX
AC AAS77840;
XX
DT 13-FEB-2002 (first entry)
XX
DE DNA encoding novel human diagnostic protein #13644.
XX
KM Human; chromosome mapping; gene mapping; gene therapy; forensic;
KM food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US008631.
XX
PR 31-MAR-2000; 2000US-00540217.
PR 23-AUG-2000; 2000US-00649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI: 2001-639362/73.
DR P-PSDB; ABG13653.
XX
PT New isolated polynucleotide and encoded polypeptides; useful in
PT diagnostics, forensics; gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX
PS Claim 1; SEQ ID NO 13644; 103pp; English.
XX
CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging

CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AS64197-AS894564 represent novel human diagnostic
CC coding sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 854 BP; 259 A; 165 C; 208 G; 222 T; 0 U; 0 Other;
Query Match 77.0%; Score 15.4; DB 5; Length 854;
Best Local Similarity 94.1%; Pred. No. 6.9e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 3 TCGCAACAGTACAATG 19
Db 833 TCGCAACAGTAAATG 817
RESULT 44
ACA20264
ID ACA20264 standard; DNA; 1134 BP.
XX
AC ACA20264;
XX
DT 19-JUN-2003 (first entry)
XX
DE Prokaryotic essential gene #1921.
XX
KM Antisense; ds; prokaryotic essential gene; cell proliferation;
KM drug design; gene.
XX
OS Staphylococcus aureus.
XX
PN WO200277183-A2.
XX
PD 03-OCT-2002.
XX
PF 21-MAR-2002; 2002WO-US009107.
XX
PR 21-MAR-2001; 2001US-00815242.
PR 06-SEP-2001; 2001US-00948993.
PR 25-OCT-2001; 2001US-0342923P.
PR 08-FEB-2002; 2002US-00072851.
PR 06-MAR-2002; 2002US-0362699P.
XX
PA (ELIT-) ELITRA PHARM INC.
XX
PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX
DR WPI: 2003-029926/02.
DR P-PSDB; ABU16394.
XX
PT New antisense nucleic acids; useful for identifying proteins or screening
PT for homologous nucleic acids required for cellular proliferation to
PT isolate candidate molecules for rational drug discovery programs.
XX
PS Claim 14; SEQ ID NO 8134; 1766pp; English.
XX
CC The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular

CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway;
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is one of the target
CC prokaryotic essential genes. Note: The sequence data for this patent did
CC not form part of the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences

CC Sequence 1134 BP; 402 A; 174 C; 230 G; 328 T; 0 U; 0 Other;

Query Match 77.0%; Score 15.4; DB 8; Length 1134;
Best Local Similarity 94.1%; Pred. No. 7.2e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AGTGGCAACAGTACAA 17
DB 389 AATGCAACAGTACAA 405

RESULT 45
AAS62439
ID AAS62439 standard; cDNA; 1704 BP.

XX AAS62439;

DT 14-FEB-2002 (first entry)

DE cDNA sequence #226 encoding novel human secreted protein.

KW Human secreted protein; hyperproliferative disorder; autoimmune disorder;

KW immune deficiency disorder; blood disorder; inflammatory disorder;

KW infectious disorder; gene therapy; antimicrobial; hepatotropic;

KW immunosuppressive; antineumatic; ss.

OS Homo sapiens.

PN WO200177291-A2.

PD 18-OCT-2001.

XX 29-MAR-2001; 2001WO-US010485.

XX 06-APR-2000; 2000US-0195604P.

XX (GENY) GENETICS INST INC.

XX Wong GG, Clark HF, Fehnel K, Agostino MJ, Howes SH, Resnick RJ;

XX Gulikota K, Graham JR;

XX WPI; 2002-010900/01.

XX New polynucleotides encoding secreted proteins useful for treating e.g.

XX asthma, HIV and Crohn's disease.

XX Claim 1; Page 192-193; 391pp; English.

XX The present invention relates to the isolation of novel cDNA sequences

XX which encode human secreted proteins. The cDNA sequences have been

CC derived from a variety of human tissues. The invention also provides a
CC method for producing proteins from these polynucleotide sequences. The
CC proteins are useful for identifying compounds that modulate their
CC activity and production, and the cell is also useful for identifying
CC compounds that modulate expression of the polynucleotide sequences
CC encoding the secreted proteins. The sequences of the invention are useful
CC for treating diseases such as hyperproliferative disorders (e.g. cancer),
CC immune deficiency disorders (e.g. severe combined immunodeficiency
CC (SCID)), autoimmune disorders (e.g. multiple sclerosis), blood disorders
CC (e.g. thrombocytopenia), inflammatory disorders (e.g. arthritis) and
CC infectious disorders (e.g. hepatitis). The polynucleotide sequences of
CC the invention are also useful in gene therapy. AAS62214-AAS62838
CC represent the cDNA sequences of the invention that encode for novel human
CC secreted proteins

CC Sequence 1704 BP; 492 A; 397 C; 379 G; 435 T; 0 U; 1 Other;

Query Match 77.0%; Score 15.4; DB 6; Length 1704;
Best Local Similarity 94.1%; Pred. No. 7.5e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 CGCAACAGTACATGG 20
DB 519 CGCAACAGTACATAG 535

RESULT 46
AAS51976
ID AAS51976 standard; DNA; 2826 BP.

XX AAS51976;

DT 13-FEB-2002 (first entry)

DE Staphylococcus aureus DNA for cellular proliferation protein #393.

KW Antisense; ds; prokaryotic cellular proliferation gene; antibiotic;

KW antibacterial; drug design.

OS Staphylococcus aureus.

PN WO200170955-A2.

PD 27-SEP-2001.

XX 21-MAR-2001; 2001WO-US009180.

XX 21-MAR-2000; 2000US-0191078P.

XX 23-MAY-2000; 2000US-0206848P.

XX 26-MAY-2000; 2000US-0207727P.

XX 23-OCT-2000; 2000US-0242578P.

XX 27-NOV-2000; 2000US-0253625P.

XX 22-DEC-2000; 2000US-0257931P.

XX 16-FEB-2001; 2001US-0263088P.

XX (ELIT) ELITRA PHARM INC.

XX Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GU;

XX Yamamoto RT, Xu HH;

XX WPI; 2001-611495/70.

XX P-PSDB; AAU34117.

XX New polynucleotides for the identification and development of
XX antibiotics, comprise sequences of antisense nucleic acids.

XX Claim 27; SEQ ID NO 458; 511pp; English.

XX The invention relates to antisense inhibitors of genes essential to

XX prokaryotic cellular proliferation, their use in identifying the genes,

XX their use in the discovery of novel antibiotics, the essential genes

XX themselves and the encoded proteins. The prokaryotes used are *Escherichia*

XX coli, *Staphylococcus aureus*, *Salmonella typhi*, *Klebsiella pneumoniae*,

CC Pseudomonas aeruginosa and Enterococcus faecalis. The invention is also
CC useful for the identification of potential new targets for antibiotic
CC development. The antisense nucleic acids can also be used to identify
CC proteins used in proliferation, to express these proteins, and to obtain
CC antibodies capable of binding to the expressed proteins. The proteins can
CC be used to screen compounds in rational drug discovery programmes. The
CC antisense nucleic acid sequence is also useful to screen for homologous
CC nucleic acids which are required for cell proliferation in a wide variety
CC of organisms. The present sequence encodes an essential prokaryotic
CC cellular proliferation protein. Note: The sequence data for this patent
CC did not form part of the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences

XX
SQ Sequence 2826 BP; 995 A; 422 C; 590 G; 819 T; 0 U; 0 Other;

Query Match 77.0%; Score 15.4; DB 4; Length 2826;
Best Local Similarity 94.1%; Pred. No. 8e+02; 1; Indels 0; Gaps 0;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AGTCGCAACAGTACAA 17
| | | | | | | | | | | | | | | | | | | | | |
389 AATCGCAACAGTACAA 405

Db

RESULT 47
ACF73815
ID ACF73815 standard; DNA; 2844 BP.
XX
XX ACF73815;
XX
DT 20-NOV-2003 (first entry)
XX
XX Staphylococcus aureus DNA #1495.
DE
XX Staphylococcus aureus DNA #1495.
XX
XX Antibacterial; vaccine; gene therapy; infection; sepsis; diagnosis;
XX enzymatic assay; antibiotic target; gene; ds.
XX
OS Staphylococcus aureus.
XX
XX WO200294868-A2.
XX
XX 28-NOV-2002.
XX
XX 27-MAR-2002; 2002WO-IB002637.
XX
XX 27-MAR-2001; 2001GB-00007661.
XX
XX (CHIR-) CHIRON SPA.
XX
XX
XX Maignani V, Mora M, Scarselli M;
XX
XX WPI; 2003-120786/11.
XX
XX P-PSDB; ABM72255.
XX
XX New Staphylococcus aureus protein, useful as a vaccine for treating or
XX preventing Staphylococcal infection, specifically an infection caused by
XX S. aureus, e.g. sepsis.
XX
XX
XX Claim 6; SEQ ID NO 2989; 49pp; English.
XX
XX The invention relates to novel genes and encoded proteins from
XX Staphylococcus aureus. A composition comprising the S. aureus protein, a
XX nucleic acid encoding the protein, or an antibody to the protein, is
XX useful as a pharmaceutical, particularly as a vaccine for treating or
XX preventing infection due to Staphylococcus bacteria, specifically an
XX infection caused by S. aureus. The composition is particularly useful for
XX treating or preventing sepsis in a patient. The composition can also be
XX used for diagnostics. The protein is also used in an assay for enzymatic
XX studies and as a target for antibiotics. This sequence represents one of
XX the novel S. aureus genes of the invention
XX
SQ Sequence 2844 BP; 1002 A; 417 C; 591 G; 834 T; 0 U; 0 Other;

Query Match 77.0%; Score 15.4; DB 8; Length 2844;
Best Local Similarity 94.1%; Pred. No. 8e+02; 1; Indels 0; Gaps 0;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AGTCGCAACAGTACAA 17
| | | | | | | | | | | | | | | | | | | | | |
389 AATCGCAACAGTACAA 405

Db

RESULT 48
AAS54776
ID AAS54776 standard; DNA; 2847 BP.
XX
XX AAS54776;
XX
DT 13-FEB-2002 (first entry)
XX
XX Staphylococcus aureus DNA for cellular proliferation protein #1088.
DE
XX Staphylococcus aureus DNA for cellular proliferation protein #1088.
XX
XX Antisense; ds; prokaryotic cellular proliferation gene; antibiotic;
XX antibacterial; drug design.
XX
XX Staphylococcus aureus.
XX
XX WO200170955-A2.
XX
XX
XX 27-SEP-2001.
XX
XX 21-MAR-2001; 2001WO-US009180.
XX
XX
XX 21-MAR-2000; 2000US-0191078P.
XX
XX 23-MAY-2000; 2000US-0206848P.
XX
XX 26-MAY-2000; 2000US-0207727P.
XX
XX 23-OCT-2000; 2000US-0242578P.
XX
XX 27-NOV-2000; 2000US-0256325P.
XX
XX 22-DEC-2000; 2000US-0257931P.
XX
XX 16-FEB-2001; 2001US-0269308P.
XX
XX
XX (ELITRA) ELITRA PHARM INC.
XX
XX
XX Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
XX Yamamoto RT, Xu HH;
XX
XX WPI; 2001-611495/70.
XX
XX P-PSDB; AAU36917.
XX
XX New polynucleotides for the identification and development of
XX antibiotics, comprise sequences of antisense nucleic acids.
XX
XX
XX Claim 27; SEQ ID NO 8413; 51pp; English.
XX
XX The invention relates to antisense inhibitors of genes essential to
XX prokaryotic cellular proliferation, their use in identifying the genes,
XX themselves and the discovery of novel antibiotics, the essential genes
XX and their use in the discovery of novel antibiotics. The prokaryotes used are Escherichia
XX coli, Staphylococcus aureus, Salmonella typhi, Klebsiella pneumoniae,
XX Pseudomonas aeruginosa and Enterococcus faecalis. The invention is also
XX useful for the identification of potential new targets for antibiotic
XX development. The antisense nucleic acids can also be used to identify
XX proteins used in proliferation, to express these proteins, and to obtain
XX antibodies capable of binding to the expressed proteins. The proteins can
XX be used to screen compounds in rational drug discovery programmes. The
XX antisense nucleic acid sequence is also useful to screen for homologous
XX nucleic acids which are required for cell proliferation in a wide variety
XX of organisms. The present sequence encodes an essential prokaryotic
XX cellular proliferation protein. Note: The sequence data for this patent
XX did not form part of the printed specification, but was obtained in
XX electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences

XX
SQ Sequence 2847 BP; 1008 A; 424 C; 592 G; 823 T; 0 U; 0 Other;

Query Match 77.0%; Score 15.4; DB 4; Length 2847;
Best Local Similarity 94.1%; Pred. No. 8e+02; 1; Indels 0; Gaps 0;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 AGTGCACAAACAGTACAA 17
Db 389 AATGCCAAACAGTACAA 405

RESULT 49

ACA20256
ID ACA20256 standard; DNA; 2847 BP.

AC ACA20256;

DT 19-JUN-2003 (first entry)

DE Prokaryotic essential gene #1913.

XM Antisense; ds; prokaryotic essential gene; cell proliferation;

XX drug design; gene.

OS Staphylococcus aureus.

PN MO200277183-A2.

PD 03-OCT-2002.

PF 21-MAR-2002; 2002WC-US009107.

PR 21-MAR-2001; 2001US-00815242.

PR 06-SEP-2001; 2001US-00948993.

PR 25-OCT-2001; 2001US-0342923P.

PR 08-FEB-2002; 2002US-00072851.

PR 06-MAR-2002; 2002US-0362699P.

XX (EITR-) ELITRA PHARM INC.

XX Wang L, Zammudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JM;

PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Foreyth RA, Xu HH;

XX WPI: 2003-029926/02.

XX P-PSDB; ABU16386.

XX New antisense nucleic acids, useful for identifying proteins or screening

PT for homologous nucleic acids required for cellular proliferation to

PT isolate candidate molecules for rational drug discovery programs.

XX Claim 14; SEQ ID NO 8126; 1766bp; English.

XX The invention relates to an isolated nucleic acid comprising any one of

CC the 6213 antisense sequences given in the specification where expression

CC of the nucleic acid inhibits proliferation of a cell. Also included are:

CC (1) a vector comprising a promoter operably linked to the nucleic acid

CC encoding a polypeptide whose expression is inhibited by the antisense

CC nucleic acid; (2) a host cell containing the vector; (3) an isolated

CC polypeptide or its fragment whose expression is inhibited by the

CC antisense nucleic acid; (4) an antibody capable of specifically binding

CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular

CC proliferation; (7) identifying a gene in an operon required for

CC proliferation; (8) identifying a compound that influences the activity of

CC the gene product or that has an activity against a biological pathway; (9)

CC required for proliferation, or that inhibits cellular proliferation; (10)

CC identifying a gene required for cellular proliferation or the biological

CC pathway in which a proliferation-required gene or its gene product lies

CC or a gene on which the test compound that inhibits proliferation of an

CC organism acts; (9) manufacturing an antibiotic; (10) profiling a

CC compound's activity; (11) a culture comprising strains in which the gene

CC product is overexpressed or underexpressed; (12) determining the extent

CC to which each of the strains is present in a culture or collection of

CC strains; or (13) identifying the target of a compound that inhibits the

CC proliferation of an organism. The antisense nucleic acids are useful for

CC identifying proteins or screening for homologous nucleic acids required

CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than S. aureus, S. typhimurium,
CC K. pneumoniae or P. aeruginosa. The present sequence is one of the target
CC prokaryotic essential genes. Note: The sequence data for this patent did
CC not form part of the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences

XX Sequence 2847 BP; 1005 A; 419 C; 592 G; 831 T; 0 U; 0 Other;

Query Match 77.0%; Score 15.4; DB 8; Length 2847;
Best Local Similarity 94.1%; Pred. No. 8e+02; 1; Indels 0; Gaps 0;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 AGTGCACAAACAGTACAA 17
Db 389 AATGCCAAACAGTACAA 405

RESULT 50
ABL29542/C
ID ABL29542 standard; DNA; 4971 BP.

XX ABL29542;

DT 26-MAR-2002 (first entry)

DE Drosophila melanogaster genomic polynucleotide SEQ ID NO 40099.

XX Drosophila; developmental biology; cell signalling; insecticide;

XX K. pharmaceutical; gene; ds.

XX Drosophila melanogaster.

XX MO200171042-A2.

XX 27-SEP-2001.

XX 23-MAR-2001; 2001WC-US009231.

XX 23-MAR-2000; 2000US-0191637P.

XX 11-JUL-2000; 2000US-00614150.

XX (PEKE) PE CORP NY.

XX Venter JC, Adams M, Li PWD, Myers EW;

XX WPI: 2001-656860/75.

XX Claim 1; SEQ ID NO 40099; 21pp + Sequence Listing; English.

XX The invention relates to an isolated nucleic acid detection reagent

CC capable of detecting 1000 or more genes from Drosophila. The invention is

CC useful in developmental biology and in elucidating cell signalling and

CC cell-cell interactions in higher eukaryotes for the development of

CC insecticides, therapeutics and pharmaceutical drugs. The invention

CC discloses genomic DNA sequences (ABL16176-ABL10511), expressed DNA

CC sequences (ABL01840-ABL16175) and the encoded proteins (ABR57737-

CC ABR57202). The sequence data for this patent did not form part of the

CC printed specification, but was obtained in electronic format directly

CC from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX Sequence 4971 BP; 1370 A; 1099 C; 1069 G; 1433 T; 0 U; 0 Other;

Query Match 77.0%; Score 15.4; DB 4; Length 4971;

Best Local Similarity 94.1%; Pred. No. 8.5e+02; 1; Indels 0; Gaps 0;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Mon Dec 6 12:24:36 2004

us-10-050-189a-7.rng

Page 31

Qy 4 CGCAACAGTCAATGG 20
|||
Db 760 CGCAACAGGCAATGG 744
|||

Search completed: December 3, 2004, 02:24:51
Job time : 321.789 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Comphen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 2, 2004, 23:31:20 ; Search time 626.579 Seconds.
(without alignments)
1509.457 Million cell updates/sec

Title: US-10-050-189a-7
Perfect score: 20
Sequence: 1 agtcgaacagctacatg 20

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 4526729 segs, 23644849745 residues
Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 1000 summaries

Database :

GenBml:.*
1: gb_ba:.*
2: gb_hlg:.*
3: gb_in:.*
4: gb_om:.*
5: gb_ov:.*
6: gb_pat:.*
7: gb_ph:.*
8: gb_pl:.*
9: gb_pr:.*
10: gb_ro:.*
11: gb_sts:.*
12: gb_sy:.*
13: gb_un:.*
14: gb_vi:.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	20	100.0	20	6	AX481360 Sequence
2	20	100.0	66479	6	AX676048 Sequence
3	20	100.0	78376	9	AL359692 Human DNA
4	17.4	87.0	110000	2	AC148981-2 Continuation (3 of
5	17.4	87.0	135276	2	AC073603 Mus muscu
6	17.4	87.0	151365	2	AC140732 Rattus no
7	17.4	87.0	202183	2	AC099737 Mus muscu
8	17.4	87.0	205681	2	AC119471 Rattus no
9	17.4	87.0	215196	2	AC073796 Mus muscu
10	17.4	87.0	229472	2	AC149052 Mus muscu
11	17.4	87.0	243210	2	AC140735 Rattus no
12	17.4	87.0	243702	2	AC136831 Rattus no
13	17.4	87.0	243702	2	AC115331 Rattus no
14	17.4	87.0	289856	2	AC094556 Rattus no
15	16.8	84.0	10601	14	HHV6U1102 X92436 Human herpe
16	16.8	84.0	56884	2	AC104588 Homo sapi
17	16.8	84.0	71049	2	AC104586 Homo sapi
18	16.8	84.0	110000	2	AC112029 Continuation (3 of
19	16.8	84.0	110000	2	AC112029_3 Continuation (4 of

20	16.8	84.0	112404	8	OSJN00288
21	16.8	84.0	116585	2	AL139153
22	16.8	84.0	120311	10	AC006945
23	16.8	84.0	128294	8	AC135225
24	16.8	84.0	137580	8	AF377947
25	16.8	84.0	155848	9	AC104996
26	16.8	84.0	159321	14	HHV6ACNM
27	16.8	84.0	159693	2	OSJN00286
28	16.8	84.0	161573	9	AL359894
29	16.8	84.0	161573	14	AB021506
30	16.8	84.0	162114	14	AF157706
31	16.8	84.0	164107	9	AC104982
32	16.8	84.0	167273	10	AC083894
33	16.8	84.0	178546	2	AC122943
34	16.8	84.0	180212	2	AC140318
35	16.8	84.0	187725	10	AC147514
36	16.8	84.0	207091	10	AC079443
37	16.8	84.0	229316	2	AC132022
38	16.8	84.0	240243	2	AC112824
39	16.8	84.0	282567	2	AC099214
40	16.8	84.0	321695	1	BX571870
41	16.8	84.0	348505	1	TRAE56651
42	16.4	82.0	925	8	CO574602
43	16.4	82.0	950	6	CO574602
44	16.4	82.0	1143	3	AY060813
45	16.4	82.0	2901	6	CO594254
46	16.4	82.0	3165	6	CO574601
47	16.4	82.0	89765	2	AC017970
48	16.4	82.0	100578	9	AL390057
49	16.4	82.0	108623	5	BX649472
50	16.4	82.0	121478	9	AL365229
51	16.4	82.0	131111	10	AC140781
52	16.4	82.0	146687	9	AC110592
53	16.4	82.0	150053	8	AP003735
54	16.4	82.0	158144	2	CR550301
55	16.4	82.0	174228	2	AC016105
56	16.4	82.0	188872	3	AC099009
57	16.4	82.0	189125	2	CR392339
58	16.4	82.0	191587	10	AC117610
59	16.4	82.0	192338	2	AC022221
60	16.4	82.0	193549	9	AC087286
61	16.4	82.0	194575	2	AC023140
62	16.4	82.0	196337	3	AC005894
63	16.4	82.0	200626	8	AC109929
64	16.4	82.0	204616	2	AC132885
65	16.4	82.0	242172	3	AB003832
66	16.4	82.0	244805	2	CR382321
67	16.4	82.0	302963	2	CR394524
68	16.4	82.0	321304	3	CEY105C58
69	16.4	80.0	112559	2	AC087120
70	16.4	80.0	177688	10	AC127274
71	16.4	80.0	187484	2	AC119772
72	16.4	80.0	210560	2	AC131578
73	16.4	80.0	245764	2	AC106238
74	16.4	80.0	261998	2	AC107722
75	16.4	80.0	282009	2	AC131579
76	15.8	79.0	182	6	AX237323
77	15.8	79.0	182	6	AX237477
78	15.8	79.0	190	3	SUR5818
79	15.8	79.0	431	6	AX778174
80	15.8	79.0	440	6	AX779741
81	15.8	79.0	642	11	PM5D3B
82	15.8	79.0	903	6	AX568554
83	15.8	79.0	906	5	AR481547
84	15.8	79.0	955	5	BX322307
85	15.8	79.0	983	5	BX34761
86	15.8	79.0	1000	8	AK065851
87	15.8	79.0	1052	8	AK119532
88	15.8	79.0	1281	6	AR375579
89	15.8	79.0	1353	5	CR338752
90	15.8	79.0	1523	10	BC024458
91	15.8	79.0	1704	6	AR318340
92	15.8	79.0	1965	3	AF495359

AL731639	Oryza sat
AL139153	Homo sapi
AC006945	Mus muscu
AC135225	Oryza sat
AF377947	Oryza sat
AC104996	Homo sapi
HHV6ACNM	X83413 Human herpe
AL731641	Oryza sat
AL359894	Human DNA
AB021506	Human her
AF157706	Human her
AC104982	Homo sapi
AC083894	Mus muscu
AC122943	Mus muscu
AC140318	Mus muscu
AC147514	Mus muscu
AC079443	Mus muscu
AC132022	Rattus no
AC112824	Rattus no
AC099214	Mus muscu
BX571870	Photograph
TRAE56651	Triticum
CO574602	Sequence
AY060813	Drosophila
CO594254	Sequence
CO574601	Sequence
AC017970	Drosophila
AL390057	Human DNA
BX649472	Zebrafish
AL365229	Human DNA
AC140781	Mus muscu
AC10592	Homo sapi
AP003735	Oryza sat
CR550301	Danio rer
AC016105	Homo sapi
AC099009	Drosophila
CR392339	Danio rer
AC117610	Mus muscu
AC022221	Homo sapi
AC087286	Homo sapi
AC023140	Homo sapi
AC005894	Drosophila
AC109929	Oryza sat
AC132885	Mus muscu
AB003832	Drosophila
CR382321	Danio rer
CR394524	Danio rer
CEY105C58	Caenorhab
AC087120	Mus muscu
AC127274	Mus muscu
AC119772	Rattus no
AC131578	Mus muscu
AC106238	Rattus no
AC107722	Mus muscu
AC131579	Mus muscu
AX237323	Sequence
AX237477	Sequence
AX778174	Sequence
AX779741	Sequence
AX568554	Sequence
AR481547	Sequence
BX322307	Gallus ga
BX34761	Gallus ga
AK065851	Oryza sat
AK119532	Oryza sat
AR375579	Sequence
CR338752	Gallus ga
BC024458	Mus muscu
AR318340	Sequence
AF495359	Drosophila

93	15.8	79.0	1966	3	AY128486	AY128486 Drosophila	166	15.8	79.0	166660	9	AC008073	AC008073 Homo sapi
94	15.8	79.0	2004	6	AX461258	AX461258 Sequence	167	15.8	79.0	168994	9	AC097480	AC097480 Homo sapi
95	15.8	79.0	2330	6	C0714512	C0714512 Sequence	168	15.8	79.0	169834	10	AC124430	AC124430 Mus muscu
96	15.8	79.0	2417	9	HSU10485	HSU10485 Human lymph	169	15.8	79.0	171089	9	AC006368	AC006368 Homo sapi
97	15.8	79.0	2716	6	AX781111	AX781111 Sequence	170	15.8	79.0	171577	10	AC106830	AC106830 Mus muscu
98	15.8	79.0	2772	6	C0576033	C0576033 Sequence	171	15.8	79.0	173519	9	AC145860	AC145860 Pan trogl
99	15.8	79.0	4183	10	RNGABAR1	RNGABAR1 Sequence	172	15.8	79.0	173519	9	AC023108	AC023108 Homo sapi
100	15.8	79.0	8876	6	CQ789079	CQ789079 Sequence	173	15.8	79.0	173956	9	AC105383	AC105383 Homo sapi
101	15.8	79.0	8876	6	AR218938	AR218938 Sequence	174	15.8	79.0	174133	8	AF083031	AF083031 Guillard
102	15.8	79.0	8876	6	BD003850	BD003850 Polymorph	175	15.8	79.0	175554	2	AC119624	AC119624 Rattus no
103	15.8	79.0	10029	1	AE007396	AE007396 Streptococ	176	15.8	79.0	175752	2	AC091721	AC091721 Pan trogl
104	15.8	79.0	10615	1	AE008456	AE008456 Streptococ	177	15.8	79.0	175808	9	AC146040	AC146040 Homo sapi
105	15.8	79.0	10846	1	AE013594	AE013594 Methanosa	178	15.8	79.0	177716	5	AP001027	AP001027 Homo sapi
106	15.8	79.0	14399	8	SPAC1F12	SPAC1F12 S. pombe chr	179	15.8	79.0	177716	5	BX539324	BX539324 Zebrafish
107	15.8	79.0	26333	3	CBRG17X04	CBRG17X04 Caenorhab	180	15.8	79.0	178403	5	AL845283	AL845283 Mouse DNA
108	15.8	79.0	26713	9	CBRG17X04	CBRG17X04 Caenorhab	181	15.8	79.0	178403	5	BX088688	BX088688 Zebrafish
109	15.8	79.0	37521	2	AC149354	AC149354 Homo sapi	182	15.8	79.0	180561	2	AP002409	AP002409 Homo sapi
110	15.8	79.0	46077	2	AC137289	AC137289 Rattus no	183	15.8	79.0	182041	2	AC136153	AC136153 Mus muscu
111	15.8	79.0	51424	5	BX276084	BX276084 Zebrafish	184	15.8	79.0	182228	10	AC036145	AC036145 Mus muscu
112	15.8	79.0	52204	2	AC015888	AC015888 Homo sapi	185	15.8	79.0	182712	9	AC018953	AC018953 Homo sapi
113	15.8	79.0	68191	2	AC136387	AC136387 Rattus no	186	15.8	79.0	183368	2	AC117878	AC117878 Rattus no
114	15.8	79.0	68548	9	HSU70984	HSU70984 Human Xp22	187	15.8	79.0	183602	2	AC024025	AC024025 Homo sapi
115	15.8	79.0	72652	8	NCB912	NCB912 Neurospor	188	15.8	79.0	184781	2	CR388151	CR388151 Danio rer
116	15.8	79.0	74417	8	NCB1D4	NCB1D4 Neurospor	189	15.8	79.0	184892	10	AC016017	AC016017 Mus muscu
117	15.8	79.0	82675	2	AC139375	AC139375 Mus muscu	190	15.8	79.0	186822	10	AL837505	AL837505 Mouse DNA
118	15.8	79.0	88366	2	AC104193	AC104193 Mus muscu	191	15.8	79.0	187972	2	AC146096	AC146096 Pan trogl
119	15.8	79.0	93225	2	AC021825	AC021825 Homo sapi	192	15.8	79.0	189458	9	AC091602	AC091602 Homo sapi
120	15.8	79.0	94038	8	AC004401	AC004401 Arabidops	193	15.8	79.0	190298	10	AC102676	AC102676 Mus muscu
121	15.8	79.0	107100	9	AC117377	AC117377 Homo sapi	194	15.8	79.0	195167	2	AC119829	AC119829 Mus muscu
122	15.8	79.0	110100	9	AC108405	AC108405 Mus muscu	195	15.8	79.0	195178	2	AC093471	AC093471 Mus muscu
123	15.8	79.0	110000	2	AC123189_4	AC123189_4 Continuation (5 of	196	15.8	79.0	195187	2	BX957308	BX957308 Danio rer
124	15.8	79.0	112229	2	AC150129	AC150129 Gallus ga	197	15.8	79.0	195316	5	AL923082	AL923082 Zebrafish
125	15.8	79.0	112679	2	AC142219	AC142219 Rattus no	198	15.8	79.0	195316	5	AL923082	AL923082 Zebrafish
126	15.8	79.0	113006	6	AC006978	AC006978 Homo sapi	199	15.8	79.0	195667	10	AC023175	AC023175 Mus muscu
127	15.8	79.0	113769	9	AP003799	AP003799 Oryza sat	200	15.8	79.0	199747	10	AL844848	AL844848 Mouse DNA
128	15.8	79.0	125937	9	AC092794	AC092794 Homo sapi	201	15.8	79.0	199908	2	AC111316	AC111316 Rattus no
129	15.8	79.0	125937	9	HS69M21	HS69M21 Human DNA	202	15.8	79.0	200108	9	AP005271	AP005271 Homo sapi
130	15.8	79.0	126315	2	AP004116	AP004116 Oryza sat	203	15.8	79.0	202742	10	AL691493	AL691493 Mouse DNA
131	15.8	79.0	127673	2	H0201G08	H0201G08 Oryza sat	204	15.8	79.0	203591	9	AC007156	AC007156 Homo sapi
132	15.8	79.0	130654	8	AP005476	AP005476 Oryza sat	205	15.8	79.0	206512	2	AC135435	AC135435 Rattus no
133	15.8	79.0	130952	8	AC146681	AC146681 Medicago	206	15.8	79.0	206877	2	AC102753	AC102753 Mus muscu
134	15.8	79.0	132017	8	AC108875	AC108875 Oryza sat	207	15.8	79.0	209021	10	AC098878	AC098878 Mus muscu
135	15.8	79.0	134426	8	AC092781	AC092781 Oryza sat	208	15.8	79.0	209561	2	CR383669	CR383669 Danio rer
136	15.8	79.0	136253	2	AC149939	AC149939 Stenomyia	209	15.8	79.0	209643	10	AC079438	AC079438 Mus muscu
137	15.8	79.0	138600	2	AC150080	AC150080 Gallus ga	210	15.8	79.0	209705	2	AC007531	AC007531 Homo sapi
138	15.8	79.0	141266	2	BX571727	BX571727 Danio rer	211	15.8	79.0	212024	2	AC073390	AC073390 Mus muscu
139	15.8	79.0	144165	5	AC008217	AC008217 Homo sapi	212	15.8	79.0	212683	2	BX936317	BX936317 Danio rer
140	15.8	79.0	144767	8	CNS08C60	AL731084 Oryza sat	213	15.8	79.0	213033	2	AC131482	AC131482 Rattus no
141	15.8	79.0	146243	2	AC150141	AC150141 Gallus ga	214	15.8	79.0	213989	5	AC145509	AC145509 Gasterost
142	15.8	79.0	146363	2	AL512423	AL512423 Homo sapi	215	15.8	79.0	216919	2	CR450833	CR450833 Danio rer
143	15.8	79.0	146415	5	AL929469	AL929469 Zebrafish	216	15.8	79.0	217249	10	AL663095	AL663095 Mouse DNA
144	15.8	79.0	146519	8	CNS08C60	AL772417 Oryza sat	217	15.8	79.0	218724	2	AC025889	AC025889 Homo sapi
145	15.8	79.0	148042	2	AC115038	AC115038 Mus muscu	218	15.8	79.0	220758	2	AC117200	AC117200 Mus muscu
146	15.8	79.0	148543	9	AC087312	AC087312 Homo sapi	219	15.8	79.0	221384	2	AC141116	AC141116 Rattus no
147	15.8	79.0	148847	2	BS000227	BS000227 Pan trogl	220	15.8	79.0	222105	2	AC142172	AC142172 Rattus no
148	15.8	79.0	149472	2	AC140281	AC140281 Mus muscu	221	15.8	79.0	224597	2	AC114874	AC114874 Rattus no
149	15.8	79.0	149970	8	AC130607	AC130607 Oryza sat	222	15.8	79.0	224846	10	AC125327	AC125327 Mus muscu
150	15.8	79.0	149992	8	BX247886	BX247886 Zebrafish	223	15.8	79.0	228999	2	AC083885	AC083885 Homo sapi
151	15.8	79.0	150682	9	AC104842	AC104842 Homo sapi	224	15.8	79.0	229465	2	AC114332	AC114332 Rattus no
152	15.8	79.0	150872	2	AC103734	AC103734 Homo sapi	225	15.8	79.0	232282	2	AC094738	AC094738 Rattus no
153	15.8	79.0	151298	5	BX537274	BX537274 Zebrafish	226	15.8	79.0	232282	2	AC103063	AC103063 Rattus no
154	15.8	79.0	151436	2	AC150177	AC150177 Gallus ga	227	15.8	79.0	235181	2	AC130063	AC130063 Rattus no
155	15.8	79.0	152890	10	AC121962	AC121962 Mus muscu	228	15.8	79.0	235946	2	AC150044	AC150044 Gallus ga
156	15.8	79.0	153241	5	AC023510	AC023510 Homo sapi	229	15.8	79.0	238803	2	AC114836	AC114836 Rattus no
157	15.8	79.0	157614	5	BX677675	BX677675 Zebrafish	230	15.8	79.0	240337	2	AC102034	AC102034 Mus muscu
158	15.8	79.0	159240	2	AC021309	AC021309 Homo sapi	231	15.8	79.0	240337	2	AC103275	AC103275 Rattus no
159	15.8	79.0	160510	10	AC124368	AC124368 Mus muscu	232	15.8	79.0	240501	2	AC110364	AC110364 Rattus no
160	15.8	79.0	161892	2	AC136667	AC136667 Rattus no	233	15.8	79.0	243046	2	AC097995	AC097995 Rattus no
161	15.8	79.0	162974	2	AC137339	AC137339 Gallus ga	234	15.8	79.0	243872	2	AC095137	AC095137 Rattus no
162	15.8	79.0	163669	2	CR339041	CR339041 Danio rer	235	15.8	79.0	246996	2	CR392340	CR392340 Danio rer
163	15.8	79.0	164819	8	AC114983	AC114983 Oryza sat	236	15.8	79.0	248254	2	SPN01903	SPN01903 Streptococ
164	15.8	79.0	165590	2	AP001076	AP001076 Homo sapi	237	15.8	79.0	249250	2	AC111494	AC111494 Rattus no
165	15.8	79.0	165590	2	AP001076	AP001076 Homo sapi	238	15.8	79.0	249250	2	AC111494	AC111494 Rattus no

C 239	15.8	79.0	253749	2	AC109560	Rattus no	C 312	15.4	77.0	132790	9	HS167A14	Z94721 Human DNA s
C 240	15.8	79.0	255836	2	AC130013	Rattus no	C 313	15.4	77.0	135726	2	CR394528	CR394528 Dantio rer
C 241	15.8	79.0	259487	2	AC124756	Mus muscu	C 314	15.4	77.0	136240	3	AC117070	AC117070 Dictyoste
C 242	15.8	79.0	260160	9	AC130454	Homo sapi	C 315	15.4	77.0	139344	2	CR392012	CR392012 Dantio rer
C 243	15.8	79.0	260193	2	AC127152	Rattus no	C 316	15.4	77.0	140881	5	BX539349	BX539349 Zebrafish
C 244	15.8	79.0	268909	2	AC112556	Rattus no	C 317	15.4	77.0	144022	5	BX089594	BX089594 Zebrafish
C 245	15.8	79.0	269096	5	BX004800	Zebrafish	C 318	15.4	77.0	144131	10	AC110822	AC110822 Mus muscu
C 246	15.8	79.0	271519	2	AC020887	Mus muscu	C 319	15.4	77.0	145536	5	BX682233	BX682233 Zebrafish
C 247	15.8	79.0	272239	2	AL928724	Dantio rer	C 320	15.4	77.0	146204	5	BX323080	BX323080 Zebrafish
C 248	15.8	79.0	274570	2	AC111994	Rattus no	C 321	15.4	77.0	146885	2	BX957322	BX957322 Dantio rer
C 249	15.8	79.0	276724	2	AC110104	Rattus no	C 322	15.4	77.0	148076	2	CR385021	CR385021 Dantio rer
C 250	15.8	79.0	278521	2	AC093715	Homo sapi	C 323	15.4	77.0	148491	2	CR626666	CR626666 Dantio rer
C 251	15.8	79.0	278781	2	AC099235	Rattus no	C 324	15.4	77.0	148904	2	BX088527	BX088527 Zebrafish
C 252	15.8	79.0	290841	1	AE017322	Listeria	C 325	15.4	77.0	149122	2	BX511502	BX511502 Dantio rer
C 253	15.8	79.0	296613	2	AC135877	Rattus no	C 326	15.4	77.0	149316	5	AL929019	AL929019 Zebrafish
C 254	15.8	79.0	341944	2	AC079158	Homo sapi	C 327	15.4	77.0	153929	2	AC011883	AC011883 Homo sapi
C 255	15.8	79.0	349980	6	AX571762	Sequence	C 328	15.4	77.0	155542	2	BX510929	BX510929 Dantio rer
C 256	15.4	77.0	306	6	AX541649	Sequence	C 329	15.4	77.0	155820	2	BX546478	BX546478 Dantio rer
C 257	15.4	77.0	506	11	BV136587	PA2A00042	C 330	15.4	77.0	156583	2	CR394535	CR394535 Dantio rer
C 258	15.4	77.0	542	6	CQ100115	Sequence	C 331	15.4	77.0	156938	5	BX510305	BX510305 Zebrafish
C 259	15.4	77.0	542	6	CQ139105	Sequence	C 332	15.4	77.0	157344	5	BX545847	BX545847 Zebrafish
C 260	15.4	77.0	542	6	CQ175700	Sequence	C 333	15.4	77.0	157379	5	AL928946	AL928946 Zebrafish
C 261	15.4	77.0	542	6	CQ222495	Sequence	C 334	15.4	77.0	158387	2	BX470205	BX470205 Dantio rer
C 262	15.4	77.0	542	6	CQ260448	Sequence	C 335	15.4	77.0	158573	5	BX248105	BX248105 Zebrafish
C 263	15.4	77.0	542	6	CQ298090	Sequence	C 336	15.4	77.0	159419	5	AC144823	AC144823 Dantio rer
C 264	15.4	77.0	542	6	CQ334579	Sequence	C 337	15.4	77.0	159928	3	AC010565	AC010565 Drosophil
C 265	15.4	77.0	573	8	AB114650	Citrus un	C 338	15.4	77.0	161362	2	CR394555	CR394555 Dantio rer
C 266	15.4	77.0	573	8	AB114658	Citrus si	C 339	15.4	77.0	161366	5	BX005382	BX005382 Zebrafish
C 267	15.4	77.0	573	8	AB114666	Citrus li	C 340	15.4	77.0	161508	2	BX927222	BX927222 Dantio rer
C 268	15.4	77.0	715	3	AB054188	Panauus 1	C 341	15.4	77.0	161518	2	BX679660	BX679660 Dantio rer
C 269	15.4	77.0	819	3	BC075123	Xenopus 1	C 342	15.4	77.0	161982	10	AC124356	AC124356 Mus muscu
C 270	15.4	77.0	965	3	AB054989	Panauus 1	C 343	15.4	77.0	162019	5	BX005271	BX005271 Zebrafish
C 271	15.4	77.0	1713	8	AB072343	Citrus un	C 344	15.4	77.0	163121	5	AL928637	AL928637 Zebrafish
C 272	15.4	77.0	1907	8	CS1319762	Citrus si	C 345	15.4	77.0	163321	2	BX571969	BX571969 Dantio rer
C 273	15.4	77.0	2245	8	AF372617	Citrus x	C 346	15.4	77.0	164093	2	CR376741	CR376741 Dantio rer
C 274	15.4	77.0	2536	3	AY061826	Drosophil	C 347	15.4	77.0	164628	9	CN601DVJ	AL358588 Human chr
C 275	15.4	77.0	2844	6	AX620026	Sequence	C 348	15.4	77.0	165022	9	AC012289	AC012289 Homo sapi
C 276	15.4	77.0	3184	5	BC060440	Xenopus 1	C 349	15.4	77.0	166396	2	AC023957	AC023957 Homo sapi
C 277	15.4	77.0	4971	6	CO613796	Sequence	C 350	15.4	77.0	166468	2	BX363658	BX363658 Dantio rer
C 278	15.4	77.0	10883	8	SPU57841	Schistosach	C 351	15.4	77.0	167187	3	AC010108	AC010108 Drosophil
C 279	15.4	77.0	10999	1	AE002536	Neisseria	C 352	15.4	77.0	167743	2	BX322580	BX322580 Dantio rer
C 280	15.4	77.0	15286	6	AF194079	Neisseria	C 353	15.4	77.0	168834	5	CR457462	CR457462 Dantio rer
C 281	15.4	77.0	15424	6	CO613766	Sequence	C 354	15.4	77.0	170665	5	BX072653	BX072653 Zebrafish
C 282	15.4	77.0	37077	5	BX927083	Zebrafish	C 355	15.4	77.0	171015	5	BX005395	BX005395 Zebrafish
C 283	15.4	77.0	37077	5	BX276119	Zebrafish	C 356	15.4	77.0	171830	5	AL772273	AL772273 Zebrafish
C 284	15.4	77.0	3210	3	U41508	Ctenorhaddi	C 357	15.4	77.0	172367	2	BX572636	BX572636 Dantio rer
C 285	15.4	77.0	53354	1	AC078935	Staphyloc	C 358	15.4	77.0	173127	9	AC021451	AC021451 Homo sapi
C 286	15.4	77.0	57214	2	AC019870	Drosophil	C 359	15.4	77.0	173219	2	BX897738	BX897738 Dantio rer
C 287	15.4	77.0	59620	2	AC014068	Drosophil	C 360	15.4	77.0	174135	9	AL359637	AL359637 Human DNA
C 288	15.4	77.0	66378	2	AC113131	Homo sapi	C 361	15.4	77.0	174257	2	BX080755	BX080755 Dantio rer
C 289	15.4	77.0	66378	2	AC001497	Drosophil	C 362	15.4	77.0	174257	2	CR405666	CR405666 Dantio rer
C 290	15.4	77.0	77731	5	AC087254	Dantio rer	C 363	15.4	77.0	174286	2	CR405666	CR405666 Dantio rer
C 291	15.4	77.0	86950	2	AC149547	Medicago	C 364	15.4	77.0	176490	9	AC079148	AC079148 Homo sapi
C 292	15.4	77.0	93853	2	CR388160	Continuati	C 365	15.4	77.0	176890	9	BX248511	BX248511 Zebrafish
C 293	15.4	77.0	95798	2	AL954721	Zebrafish	C 366	15.4	77.0	176986	5	BX48511	BX48511 Zebrafish
C 294	15.4	77.0	97774	5	AL591399	Zebrafish	C 367	15.4	77.0	177080	5	BX323984	BX323984 Zebrafish
C 295	15.4	77.0	99681	5	AL713850	Zebrafish	C 368	15.4	77.0	177608	3	AC099553	AC099553 Drosophil
C 296	15.4	77.0	105510	5	BX005456	Zebrafish	C 369	15.4	77.0	178874	2	BX323590	BX323590 Dantio rer
C 297	15.4	77.0	105510	5	AC149582	Zebrafish	C 370	15.4	77.0	179005	9	AC008507	AC008507 Homo sapi
C 298	15.4	77.0	105607	5	BX247884	Zebrafish	C 371	15.4	77.0	180137	2	AC023081	AC023081 Homo sapi
C 299	15.4	77.0	108530	2	BX571856	Continuati	C 372	15.4	77.0	180305	5	BX004768	BX004768 Zebrafish
C 300	15.4	77.0	110000	1	BX571857	Continuati	C 373	15.4	77.0	180835	2	BX957237	BX957237 Dantio rer
C 301	15.4	77.0	110000	1	BX571857	Continuati	C 374	15.4	77.0	180835	2	CR391970	CR391970 Dantio rer
C 302	15.4	77.0	110000	2	BX321891	Continuati	C 375	15.4	77.0	181320	5	BX547992	BX547992 Zebrafish
C 303	15.4	77.0	110000	2	BX321891	Continuati	C 376	15.4	77.0	182960	5	AC010017	AC010017 Drosophil
C 304	15.4	77.0	110000	2	BX321891	Continuati	C 377	15.4	77.0	183026	5	BX323086	BX323086 Zebrafish
C 305	15.4	77.0	110000	2	BX908723	Continuati	C 378	15.4	77.0	183715	2	CR626906	CR626906 Dantio rer
C 306	15.4	77.0	110000	2	BX908723	Continuati	C 379	15.4	77.0	184431	2	CR628331	CR628331 Dantio rer
C 307	15.4	77.0	114149	5	BX470242	Zebrafish	C 380	15.4	77.0	185515	2	BX957362	BX957362 Dantio rer
C 308	15.4	77.0	115248	5	BX255508	Zebrafish	C 381	15.4	77.0	186780	2	CR626943	CR626943 Dantio rer
C 309	15.4	77.0	124224	5	BX294657	Zebrafish	C 382	15.4	77.0	187061	2	AC122564	AC122564 Mus muscu
C 310	15.4	77.0	128110	5	BX936334	Dantio rer	C 383	15.4	77.0	188341	3	AC005719	AC005719 Drosophil
C 311	15.4	77.0	128110	2	BX936334	Dantio rer	C 384	15.4	77.0	188529	2	BX927370	BX927370 Dantio rer

C 385	15.4	77.0	188746	5	AL929594	AL929594 Zebrafish	C 458	15.4	77.0	349980	6	AX417043	AX417043 Sequence
C 386	15.4	77.0	190095	2	BX276112	BX276112 Danio rer	C 459	15.2	76.0	174	4	AY530078	AY530078 Solenodon
C 387	15.4	77.0	191663	2	BX649269	BX649269 Danio rer	C 460	15.2	76.0	245	11	BV080588	BV080588 sc1218_p6
C 388	15.4	77.0	193472	2	CR388077	CR388077 Danio rer	C 461	15.2	76.0	245	11	BV080591	BV080591 sc1218_p6
C 389	15.4	77.0	193505	2	CR628366	CR628366 Danio rer	C 462	15.2	76.0	246	11	BV080580	BV080580 sc1218_p5
C 390	15.4	77.0	195651	5	BX294114	BX294114 Zebrafish	C 463	15.2	76.0	246	11	BV080582	BV080582 sc1218_p6
C 391	15.4	77.0	198070	2	BX649326	BX649326 Danio rer	C 464	15.2	76.0	246	11	BV080586	BV080586 sc1218_p6
C 392	15.4	77.0	199540	10	AL985952	AL985952 Mouse DNA	C 465	15.2	76.0	246	11	BV080589	BV080589 sc1218_p6
C 393	15.4	77.0	200006	2	BX004997	BX004997 Danio rer	C 466	15.2	76.0	246	11	BV080590	BV080590 sc1218_p6
C 394	15.4	77.0	200937	2	BX510646	BX510646 Danio rer	C 467	15.2	76.0	246	11	BV080592	BV080592 sc1218_p6
C 395	15.4	77.0	201132	2	BX927146	BX927146 Danio rer	C 468	15.2	76.0	246	11	BV080593	BV080593 sc1218_p6
C 396	15.4	77.0	201188	5	AC145974	AC145974 Gallus ga	C 469	15.2	76.0	254	11	BV080584	BV080584 sc1218_p6
C 397	15.4	77.0	201816	5	BX248121	BX248121 Zebrafish	C 470	15.2	76.0	254	11	BV080585	BV080585 sc1218_p6
C 398	15.4	77.0	203371	2	AC139623	AC139623 Danio rer	C 471	15.2	76.0	254	11	BV080587	BV080587 sc1218_p6
C 399	15.4	77.0	203944	2	BX537255	BX537255 Danio rer	C 472	15.2	76.0	342	6	AX918716	AX918716 Sequence
C 400	15.4	77.0	205102	2	CR336225	CR336225 Danio rer	C 473	15.2	76.0	342	6	BD054249	BD054249 Sequence
C 401	15.4	77.0	205289	2	AL954337	AL954337 Zebrafish	C 474	15.2	76.0	344	11	BV108269	BV108269 P2A01791
C 402	15.4	77.0	205802	5	AL772132	AL772132 Zebrafish	C 475	15.2	76.0	347	11	BV108277	BV108277 P2A01791
C 403	15.4	77.0	207146	2	CR388142	CR388142 Danio rer	C 476	15.2	76.0	349	11	BV108265	BV108265 P2A01791
C 404	15.4	77.0	207701	2	BX545856	BX545856 Danio rer	C 477	15.2	76.0	357	4	AF271977	AF271977 Bos tauru
C 405	15.4	77.0	208075	2	BX957363	BX957363 Danio rer	C 478	15.2	76.0	368	11	BV108271	BV108271 P2A01791
C 406	15.4	77.0	208770	2	CR388135	CR388135 Danio rer	C 479	15.2	76.0	372	11	BV108274	BV108274 P2A01791
C 407	15.4	77.0	209782	10	AL603905	AL603905 Mouse DNA	C 480	15.2	76.0	378	11	BV108272	BV108272 P2A01791
C 408	15.4	77.0	210355	2	BX511081	BX511081 Danio rer	C 481	15.2	76.0	380	11	BV108276	BV108276 P2A01791
C 409	15.4	77.0	210630	5	AL929276	AL929276 Zebrafish	C 482	15.2	76.0	384	11	BV108270	BV108270 P2A01791
C 410	15.4	77.0	212049	5	BX005484	BX005484 Zebrafish	C 483	15.2	76.0	385	11	BV108264	BV108264 P2A01791
C 411	15.4	77.0	212160	5	BX323453	BX323453 Zebrafish	C 484	15.2	76.0	385	11	BV108266	BV108266 P2A01791
C 412	15.4	77.0	212553	2	AC099581	AC099581 Mus muscu	C 485	15.2	76.0	385	11	BV108267	BV108267 P2A01791
C 413	15.4	77.0	216389	2	BX640518	BX640518 Danio rer	C 486	15.2	76.0	385	11	BV108268	BV108268 P2A01791
C 414	15.4	77.0	216615	10	AC107850	AC107850 Mus muscu	C 487	15.2	76.0	393	11	BV108275	BV108275 P2A01791
C 415	15.4	77.0	217582	5	BX005022	BX005022 Zebrafish	C 488	15.2	76.0	396	5	AY582107	AY582107 Leptocott
C 416	15.4	77.0	222751	2	CR538727	CR538727 Danio rer	C 489	15.2	76.0	513	6	AX818942	AX818942 Sequence
C 417	15.4	77.0	223880	2	AC103105	AC103105 Rattus no	C 490	15.2	76.0	513	6	AX82972	AX82972 Sequence
C 418	15.4	77.0	224218	2	CR556716	CR556716 Danio rer	C 491	15.2	76.0	513	6	AY557727	AY557727 Saccharom
C 419	15.4	77.0	225257	2	BX088574	BX088574 Danio rer	C 492	15.2	76.0	584	11	BV038129	BV038129 S212P6970
C 420	15.4	77.0	227705	5	BX571720	BX571720 Zebrafish	C 493	15.2	76.0	756	6	AR319714	AR319714 Sequence
C 421	15.4	77.0	227892	5	BX000471	BX000471 Zebrafish	C 494	15.2	76.0	883	8	BT013102	BT013102 Lycopersi
C 422	15.4	77.0	228584	2	CR392040	CR392040 Danio rer	C 495	15.2	76.0	1010	3	AF044817	AF044817 Papilio p
C 423	15.4	77.0	228902	2	AC098116	AC098116 Rattus no	C 496	15.2	76.0	1010	3	AF044819	AF044819 Papilio m
C 424	15.4	77.0	231484	2	AC096379	AC096379 Rattus no	C 497	15.2	76.0	1010	3	AF044820	AF044820 Papilio h
C 425	15.4	77.0	233451	2	AC103060	AC103060 Rattus no	C 498	15.2	76.0	1010	3	AF044830	AF044830 Papilio c
C 426	15.4	77.0	233058	5	BX640463	BX640463 Zebrafish	C 499	15.2	76.0	1045	3	AY457621	AY457621 Papilio m
C 427	15.4	77.0	234274	2	AC148857	AC148857 Oryzomur	C 500	15.2	76.0	1121	10	BC048519	BC048519 Mus muscu
C 428	15.4	77.0	235551	2	AC130518	AC130518 Rattus no	C 501	15.2	76.0	1220	3	AY457606	AY457606 Papilio c
C 429	15.4	77.0	235897	2	AC095685	AC095685 Rattus no	C 502	15.2	76.0	1240	3	AF044828	AF044828 Papilio m
C 430	15.4	77.0	238499	5	BX005043	BX005043 Zebrafish	C 503	15.2	76.0	1276	5	CR523553	CR523553 Gallus ga
C 431	15.4	77.0	238646	2	AC099463	AC099463 Rattus no	C 504	15.2	76.0	1278	6	AX848366	AX848366 Sequence
C 432	15.4	77.0	240539	2	BX571803	BX571803 Danio rer	C 505	15.2	76.0	1349	6	AX141387	AX141387 Sequence
C 433	15.4	77.0	240539	2	BX571803	BX571803 Danio rer	C 506	15.2	76.0	1349	6	AX685942	AX685942 Sequence
C 434	15.4	77.0	240834	2	BX085193	BX085193 Danio rer	C 507	15.2	76.0	1643	5	AX934056	AX934056 Gallus ga
C 435	15.4	77.0	243235	2	CR318614	CR318614 Danio rer	C 508	15.2	76.0	1702	3	AF248864	AF248864 Stronylo
C 436	15.4	77.0	246998	2	CR450707	CR450707 Danio rer	C 509	15.2	76.0	1905	3	AT047577	AT047577 Drosophi
C 437	15.4	77.0	248738	5	BX284112	BX284112 Danio rer	C 510	15.2	76.0	1924	3	AF215861	AF215861 Drosophi
C 438	15.4	77.0	249278	5	BX511172	BX511172 Zebrafish	C 511	15.2	76.0	1941	3	MSP287537	MSP287537 Mesometra
C 439	15.4	77.0	257502	5	BX005342	BX005342 Zebrafish	C 512	15.2	76.0	1941	3	AY087725	AY087725 Arabidops
C 440	15.4	77.0	258640	2	CR382325	CR382325 Danio rer	C 513	15.2	76.0	1950	8	AY087725	AY087725 Arabidops
C 441	15.4	77.0	259889	2	AC125704	AC125704 Rattus no	C 514	15.2	76.0	2003	5	AF369901	AF369901 Xenopus l
C 442	15.4	77.0	260050	1	AL596166	AL596166 Listeria	C 515	15.2	76.0	2077	10	BC048839	BC048839 Mus muscu
C 443	15.4	77.0	261272	2	CR388059	CR388059 Danio rer	C 516	15.2	76.0	2089	5	BC048038	BC048038 Danio rer
C 444	15.4	77.0	266281	2	AE003664	AE003664 Drosophi	C 517	15.2	76.0	2248	5	AF051787	AF051787 Glycydus
C 445	15.4	77.0	266691	2	BX571847	BX571847 Danio rer	C 518	15.2	76.0	2334	6	AX934444	AX934444 Sequence
C 446	15.4	77.0	267191	2	AC132043	AC132043 Rattus no	C 519	15.2	76.0	2385	6	CQ594468	CQ594468 Sequence
C 447	15.4	77.0	282115	3	AE003545	AE003545 Drosophi	C 520	15.2	76.0	2525	10	BC057991	BC057991 Mus muscu
C 448	15.4	77.0	284624	2	BX511034	BX511034 Zebrafish	C 521	15.2	76.0	2609	3	AY548906	AY548906 Antonomsp
C 449	15.4	77.0	286150	1	AP008824	AP008824 Staphyloc	C 522	15.2	76.0	2699	9	HSTPR378	HSTPR378 H.sapiens T
C 450	15.4	77.0	290150	1	BX545912	BX545912 Danio rer	C 523	15.2	76.0	2734	10	BC067200	BC067200 Mus muscu
C 451	15.4	77.0	299050	1	AP003131	AP003131 Staphyloc	C 524	15.2	76.0	2869	1	AF290090	AF290090 Staphyloc
C 452	15.4	77.0	300409	1	AE016755	AE016755 Escherich	C 525	15.2	76.0	2955	10	BC050127	BC050127 Mus muscu
C 453	15.4	77.0	314906	3	AE003594	AE003594 Drosophi	C 526	15.2	76.0	3007	2	AC011213	AC011213 Drosophi
C 454	15.4	77.0	326301	1	NMA622491	NMA622491 Neisseria	C 527	15.2	76.0	3044	10	AY092766	AY092766 Mus muscu
C 455	15.4	77.0	326877	1	AP003360	AP003360 Staphyloc	C 528	15.2	76.0	3172	1	AF269836	AF269836 Staphyloc
C 456	15.4	77.0	349980	6	AX044034	AX044034 Sequence	C 529	15.2	76.0	3172	1	AF270274	AF270274 Staphyloc
C 457	15.4	77.0	349980	6	AX417039	AX417039 Sequence	C 530	15.2	76.0	3172	6	AR485790	AR485790 Sequence

531	15.2	76.0	3172	6	AR468228	604	15.2	76.0	95637	8	EX842594	EX842594	Neutropor
532	15.2	76.0	3172	6	AX145154	605	15.2	76.0	98613	9	AC139630	AC139630	Takifugu
533	15.2	76.0	3172	6	AX145352	606	15.2	76.0	103041	9	AC093384	AC093384	Homo sapi
534	15.2	76.0	3138	8	AF302666	607	15.2	76.0	103245	2	AC108382	AC108382	Pan trogl
535	15.2	76.0	3586	8	SCU07938	608	15.2	76.0	107113	2	AC142559	AC142559	Takifugu
536	15.2	76.0	3827	8	SCYNL243W	609	15.2	76.0	108688	8	AC105733	AC105733	Oryza sat
537	15.2	76.0	4206	3	DVZFP	610	15.2	76.0	110000	1	CR543861.22	CR543861.22	Continuation (25 o
538	15.2	76.0	4530	10	AK129249	611	15.2	76.0	110000	1	CR543861.24	CR543861.24	Continuation (25 o
539	15.2	76.0	4641	8	YSCOMP26	612	15.2	76.0	110000	2	AC091338.2	AC091338.2	Continuation (3 of
540	15.2	76.0	5130	6	CQ594467	613	15.2	76.0	110000	2	AC117128.1	AC117128.1	Continuation (2 of
541	15.2	76.0	5133	6	CQ594467	614	15.2	76.0	110000	2	AC129112.0	AC129112.0	Rattus no
542	15.2	76.0	5323	5	AB074891	615	15.2	76.0	110000	2	AC138524.1	AC138524.1	Continuation (2 of
543	15.2	76.0	5751	3	AF248863	616	15.2	76.0	110000	2	AL137126.0	AL137126.0	Al137126.0 Homo sapi
544	15.2	76.0	5919	8	SCYNL242W	617	15.2	76.0	110000	2	AL451003.0	AL451003.0	Continuation (3 of
545	15.2	76.0	5966	14	SYVCP	618	15.2	76.0	110000	2	AL583837.2	AL583837.2	Continuation (2 of
546	15.2	76.0	6187	8	SLA310657	619	15.2	76.0	110000	2	AL713976.1	AL713976.1	Continuation (11 o
547	15.2	76.0	8382	6	CQ575612	620	15.2	76.0	110000	2	AL929091.10	AL929091.10	Continuation (2 of
548	15.2	76.0	10029	1	AE013415	621	15.2	76.0	110000	2	BR294366.1	BR294366.1	Continuation (2 of
549	15.2	76.0	11021	1	AE007252	622	15.2	76.0	110000	2	BR294366.2	BR294366.2	Continuation (3 of
550	15.2	76.0	11124	1	AE009714	623	15.2	76.0	110000	8	CR382132.27	CR382132.27	Continuation (28 o
551	15.2	76.0	11565	1	AE010682	624	15.2	76.0	110000	8	CR382132.20	CR382132.20	Continuation (21 o
552	15.2	76.0	12311	8	AF004559	625	15.2	76.0	110000	9	HS46345.1	HS46345.1	Continuation (2 of
553	15.2	76.0	12461	1	AE015834	626	15.2	76.0	110000	10	AE014174.0	AE014174.0	Continuation (2 of
554	15.2	76.0	12489	1	AE014544	627	15.2	76.0	111275	2	AC139876	AC139876	AC139876 Rattus no
555	15.2	76.0	12500	1	AE001261	628	15.2	76.0	111375	2	AC073878	AC073878	Homo sapi
556	15.2	76.0	12518	1	AE015202	629	15.2	76.0	115174	9	AC106028	AC106028	Homo sapi
557	15.2	76.0	13123	1	AE003872	630	15.2	76.0	117695	2	AC137831	AC137831	Medicago
558	15.2	76.0	13852	10	AB026490	631	15.2	76.0	119504	5	BX248306	BX248306	Continuation (2 of
559	15.2	76.0	15280	4	CEH31B20	632	15.2	76.0	120466	9	AL449104	AL449104	Continuation (2 of
560	15.2	76.0	16829	4	RUMTGENOY	633	15.2	76.0	124587	2	AC141185	AC141185	Continuation (2 of
561	15.2	76.0	18155	6	AX429460	634	15.2	76.0	125150	2	AC005450	AC005450	Continuation (2 of
562	15.2	76.0	22062	6	AC019898	635	15.2	76.0	125803	8	ATP22K18	ATP22K18	Continuation (2 of
563	15.2	76.0	22253	8	SC9320X	636	15.2	76.0	126039	9	AC012665	AC012665	Continuation (2 of
564	15.2	76.0	23947	1	AE008776	637	15.2	76.0	127347	10	AC102818	AC102818	Continuation (2 of
565	15.2	76.0	24000	8	SC9320A	638	15.2	76.0	127395	7	AY176327	AY176327	Continuation (2 of
566	15.2	76.0	26568	9	HSV755G1	639	15.2	76.0	128809	9	AC021648	AC021648	Continuation (2 of
567	15.2	76.0	28534	9	EX572102	640	15.2	76.0	129800	9	AL357353	AL357353	Continuation (2 of
568	15.2	76.0	31180	3	DMC155E2	641	15.2	76.0	130325	9	AC099566	AC099566	Continuation (2 of
569	15.2	76.0	32158	3	CEW04G3	642	15.2	76.0	131815	9	AL627442	AL627442	Continuation (2 of
570	15.2	76.0	32363	9	AC114792	643	15.2	76.0	131891	9	AC008131	AC008131	Continuation (2 of
571	15.2	76.0	33171	9	AC146820	644	15.2	76.0	132675	5	BX324182	BX324182	Continuation (2 of
572	15.2	76.0	33274	3	CEFS2B5	645	15.2	76.0	132891	2	EX088564	EX088564	Continuation (2 of
573	15.2	76.0	33855	8	SCCXV33X	646	15.2	76.0	132891	2	AC002403	AC002403	Continuation (2 of
574	15.2	76.0	39800	2	AL513223	647	15.2	76.0	133915	9	AC124055	AC124055	Continuation (2 of
575	15.2	76.0	40280	3	CEBRI.7N15	648	15.2	76.0	134157	5	EX005064	EX005064	Continuation (2 of
576	15.2	76.0	40438	8	SPU4.1410	649	15.2	76.0	134325	2	AC141344	AC141344	Continuation (2 of
577	15.2	76.0	40478	3	AC068077	650	15.2	76.0	134388	2	AC141344	AC141344	Continuation (2 of
578	15.2	76.0	40482	3	U29614	651	15.2	76.0	134708	9	AC137697	AC137697	Continuation (2 of
579	15.2	76.0	40545	8	AL500524	652	15.2	76.0	134821	8	AL136092	AL136092	Continuation (2 of
580	15.2	76.0	40810	8	SPBC1773	653	15.2	76.0	135245	8	AP002485	AP002485	Continuation (2 of
581	15.2	76.0	46279	2	AC014815	654	15.2	76.0	136150	8	AP002485	AP002485	Continuation (2 of
582	15.2	76.0	49122	2	AC108385	655	15.2	76.0	136515	9	AL499609	AL499609	Continuation (2 of
583	15.2	76.0	50089	3	AC005641	656	15.2	76.0	136531	9	HS1050C22	HS1050C22	Continuation (2 of
584	15.2	76.0	51811	2	AL100162	657	15.2	76.0	137332	8	AP002820	AP002820	Continuation (2 of
585	15.2	76.0	54625	9	AL159979	658	15.2	76.0	137462	8	AP002538	AP002538	Continuation (2 of
586	15.2	76.0	58320	2	AC133850	659	15.2	76.0	138115	2	AC142040	AC142040	Continuation (2 of
587	15.2	76.0	59492	2	EX908805	660	15.2	76.0	140028	9	AC104658	AC104658	Continuation (2 of
588	15.2	76.0	64227	2	AC100013	661	15.2	76.0	140356	9	AC007528	AC007528	Continuation (2 of
589	15.2	76.0	64430	2	AC083772	662	15.2	76.0	140619	9	AC007285	AC007285	Continuation (2 of
590	15.2	76.0	67435	2	AC010479	663	15.2	76.0	141396	2	AC091709	AC091709	Continuation (2 of
591	15.2	76.0	72354	9	AL357123	664	15.2	76.0	141973	2	AC099530	AC099530	Continuation (2 of
592	15.2	76.0	73198	9	HSBJ12713	665	15.2	76.0	144062	5	AL954190	AL954190	Continuation (2 of
593	15.2	76.0	76236	2	AC023248	666	15.2	76.0	145704	3	AE002656	AE002656	Continuation (2 of
594	15.2	76.0	80236	5	AL355138	667	15.2	76.0	145880	9	HS302D9	HS302D9	Continuation (2 of
595	15.2	76.0	80804	10	EX510908	668	15.2	76.0	147851	2	AC141379	AC141379	Continuation (2 of
596	15.2	76.0	81874	5	EX323864	669	15.2	76.0	147875	9	AC067982	AC067982	Continuation (2 of
597	15.2	76.0	82291	5	AC005376	670	15.2	76.0	148819	9	AC006996	AC006996	Continuation (2 of
598	15.2	76.0	86564	1	AE008919	671	15.2	76.0	149200	9	AL159645	AL159645	Continuation (2 of
599	15.2	76.0	86966	1	AC108720	672	15.2	76.0	149427	2	AC121385	AC121385	Continuation (2 of
600	15.2	76.0	87428	3	CEY64G10A	673	15.2	76.0	149795	10	AC125173	AC125173	Continuation (2 of
601	15.2	76.0	87476	2	AC138695	674	15.2	76.0	150894	1	D90904	D90904	Continuation (2 of
602	15.2	76.0	89017	2	AC108387	675	15.2	76.0	151405	8	AP004261	AP004261	Continuation (2 of
603	15.2	76.0	91552	6	AX429455	676	15.2	76.0	151640	10	AC122053	AC122053	Continuation (2 of

677	15.2	76.0	151748	9	AC147408	Pan trogl	750	15.2	76.0	173518	10	AL772358	AL772358 Mouse DNA
678	15.2	76.0	151992	2	AC009939	Homo sapi	751	15.2	76.0	173558	3	AC093500	AC093500 Drosophila
679	15.2	76.0	152653	2	AC139129	Mus muscu	752	15.2	76.0	174223	2	AC128776	AC128776 Rattus no
680	15.2	76.0	152914	8	CNS08CAR	AL837528	753	15.2	76.0	174285	3	AC008259	AC008259 Drosophila
681	15.2	76.0	153053	2	AC119281	Mus muscu	754	15.2	76.0	175485	9	AC147539	AC147539 Pan trogl
682	15.2	76.0	153094	9	CNS05TCH	AL355098	755	15.2	76.0	175939	2	AC021191	AC021191 Homo sapi
683	15.2	76.0	153339	2	AC007350	Homo sapi	756	15.2	76.0	176759	2	AC037464	AC037464 Homo sapi
684	15.2	76.0	155415	2	AC108886	Rattus no	757	15.2	76.0	176798	2	AC149140	AC149140 Xenopus t
685	15.2	76.0	155540	5	BX324162	Zebrafish	758	15.2	76.0	177888	2	AC131687	AC131687 Mus muscu
686	15.2	76.0	155754	2	CR556723	CR556723	759	15.2	76.0	178255	9	AC019197	AC019197 Homo sapi
687	15.2	76.0	155757	10	AC112157	AC112157	760	15.2	76.0	178272	10	AC118772	AC118772 Rattus no
688	15.2	76.0	155864	2	AC120103	Rattus no	761	15.2	76.0	178477	2	AC022801	AC022801 Homo sapi
689	15.2	76.0	156647	8	AC137930	AC137930	762	15.2	76.0	178797	2	AC121048	AC121048 Rattus no
690	15.2	76.0	157666	9	AC006989	Homo sapi	763	15.2	76.0	179526	2	HS51219	AL445624 Homo sapi
691	15.2	76.0	158231	9	AC142296	AC142296	764	15.2	76.0	180313	2	AC146158	AC146158 Pan trogl
692	15.2	76.0	158828	9	AL590069	AL590069	765	15.2	76.0	180335	2	AC146276	AC146276 Par trogl
693	15.2	76.0	159089	2	CR589884	Danio rer	766	15.2	76.0	181097	2	AC118205	AC118205 Mus muscu
694	15.2	76.0	159512	9	AC145986	AC145986	767	15.2	76.0	181254	2	AC147898	AC147898 Xenopus t
695	15.2	76.0	159964	2	AC125018	Mus trogl	768	15.2	76.0	181514	2	AC147450	AC147450 Otlemur
696	15.2	76.0	160219	9	AC146277	AC146277	769	15.2	76.0	181535	2	AL592203	AL592203 Homo sapi
697	15.2	76.0	160356	9	AL356276	Human DNA	770	15.2	76.0	181859	2	AC150406	AC150406 Branchios
698	15.2	76.0	160903	9	AC092812	AC092812	771	15.2	76.0	181917	2	AC133728	AC133728 Rattus no
699	15.2	76.0	160982	2	AC128945	AC128945	772	15.2	76.0	182486	8	AC134044	AC134044 Oryza sat
700	15.2	76.0	161155	9	AP017104	AP017104	773	15.2	76.0	182666	2	AC135682	AC135682 Rattus no
701	15.2	76.0	161248	9	AC009158	AC009158	774	15.2	76.0	182859	2	AL845341	AL845341 Mus muscu
702	15.2	76.0	161393	10	AL645966	AL645966	775	15.2	76.0	183136	2	AC041000	AC041000 Homo sapi
703	15.2	76.0	161443	9	AC009884	AC009884	776	15.2	76.0	183996	9	AF196779	AF196779 Homo sapi
704	15.2	76.0	161569	3	AC104145	AC104145	777	15.2	76.0	184079	3	AC008258	AC008258 Drosophila
705	15.2	76.0	161590	9	AC016494	AC016494	778	15.2	76.0	184407	2	AC122767	AC122767 Mus muscu
706	15.2	76.0	161706	2	AC021194	AC021194	779	15.2	76.0	185170	5	BX005269	BX005269 Zebrafish
707	15.2	76.0	162224	9	AC012176	AC012176	780	15.2	76.0	185190	2	BX005148	BX005148 Danio rer
708	15.2	76.0	162689	2	AC117145	AC117145	781	15.2	76.0	185763	3	AF152363	AF152363 Homo sapi
709	15.2	76.0	162912	9	AC002067	AC002067	782	15.2	76.0	185800	2	AC079366	AC079366 Mus muscu
710	15.2	76.0	162912	10	AL627076	AL627076	783	15.2	76.0	185835	2	AC092007	AC092007 Bos tauri
711	15.2	76.0	163353	10	AC123738	AC123738	784	15.2	76.0	186331	9	AC006511	AC006511 Homo sapi
712	15.2	76.0	163762	4	AC090445	AC090445	785	15.2	76.0	186464	2	AC142067	AC142067 Rattus no
713	15.2	76.0	163894	2	AC150398	AC150398	786	15.2	76.0	186654	2	AP001772	AP001772 Homo sapi
714	15.2	76.0	165077	10	AC084382	AC084382	787	15.2	76.0	186798	10	AC140980	AC140980 Mus muscu
715	15.2	76.0	165240	2	AC011040	AC011040	788	15.2	76.0	186875	9	AL731540	AL731540 Human DNA
716	15.2	76.0	165603	9	BX548250	Zebrafish	789	15.2	76.0	186875	9	AL731540	AL731540 Human DNA
717	15.2	76.0	166686	9	AC090450	AC090450	790	15.2	76.0	187294	2	AC103917	AC103917 Drosophila
718	15.2	76.0	167015	2	BX649587	Mus muscu	791	15.2	76.0	187421	2	AC149882	AC149882 Xenopus t
719	15.2	76.0	167040	2	AC092194	AC092194	792	15.2	76.0	187941	2	AC108858	AC108858 Mus muscu
720	15.2	76.0	167543	9	AC092554	AC092554	793	15.2	76.0	188265	9	AC146204	AC146204 Pan trogl
721	15.2	76.0	167693	2	BX640399	Danio rer	794	15.2	76.0	188380	10	AC134831	AC134831 Mus muscu
722	15.2	76.0	167755	2	BX957324	Danio rer	795	15.2	76.0	188970	2	AC115749	AC115749 Mus muscu
723	15.2	76.0	167949	9	AC087821	Homo sapi	796	15.2	76.0	189953	9	AL445435	AL445435 Human DNA
724	15.2	76.0	168059	3	AC010027	AC010027	797	15.2	76.0	190094	2	AC131819	AC131819 Rattus no
725	15.2	76.0	168459	9	AL772289	AL772289	798	15.2	76.0	190454	9	AC117984	AC117984 Homo sapi
726	15.2	76.0	168486	9	CNS05TCH	AL163872	799	15.2	76.0	190647	10	AL627077	AL627077 Mouse DNA
727	15.2	76.0	168560	8	AC118132	AC118132	800	15.2	76.0	191342	3	AC104140	AC104140 Drosophila
728	15.2	76.0	168575	6	AR216265	AR216265	801	15.2	76.0	191656	2	AC118574	AC118574 Lemur cat
729	15.2	76.0	168575	6	AX134178	Sequence	802	15.2	76.0	192010	10	AL732320	AL732320 Mouse DNA
730	15.2	76.0	168575	9	AY062031	Homo sapi	803	15.2	76.0	192429	2	AC113955	AC113955 Mus muscu
731	15.2	76.0	168764	2	AC135451	AC135451	804	15.2	76.0	192768	9	AC016831	AC016831 Homo sapi
732	15.2	76.0	168838	2	AC024929	AC024929	805	15.2	76.0	192971	10	AC132367	AC132367 Mus muscu
733	15.2	76.0	169216	2	AC090428	Homo sapi	806	15.2	76.0	193404	9	AC103693	AC103693 Homo sapi
734	15.2	76.0	169434	2	AC150421	Branchios	807	15.2	76.0	193511	2	AC125907	AC125907 Rattus no
735	15.2	76.0	169758	2	BX511118	BX511118	808	15.2	76.0	193514	2	AC121674	AC121674 Rattus no
736	15.2	76.0	169876	10	AL713854	AL713854	809	15.2	76.0	193813	2	AC100723	AC100723 Mus muscu
737	15.2	76.0	170085	2	AC118718	Mus muscu	810	15.2	76.0	194935	10	AL627103	AL627103 Mouse DNA
738	15.2	76.0	170695	2	AC068545	Homo sapi	811	15.2	76.0	195205	2	AC131416	AC131416 Rattus no
739	15.2	76.0	170847	2	AC144771	AC144771	812	15.2	76.0	195428	10	AC098884	AC098884 Mus muscu
740	15.2	76.0	170892	5	CNS07EES	AL445885	813	15.2	76.0	196541	2	AC091622	AC091622 Pan trogl
741	15.2	76.0	170971	5	BX572629	Human chr	814	15.2	76.0	196797	10	AC098734	AC098734 Mus muscu
742	15.2	76.0	171215	2	BX950173	BX950173	815	15.2	76.0	196797	9	AC084866	AC084866 Homo sapi
743	15.2	76.0	171612	9	AC099742	Danio rer	816	15.2	76.0	196908	2	AC098366	AC098366 Rattus no
744	15.2	76.0	171734	10	AL954713	AL954713	817	15.2	76.0	199309	2	AC128119	AC128119 Rattus no
745	15.2	76.0	171744	9	AP001786	AP001786	818	15.2	76.0	198402	8	ATC87161	ATC87161 Homo sapi
746	15.2	76.0	172409	5	BX649268	Zebrafish	819	15.2	76.0	199184	2	AC023272	AC023272 Homo sapi
747	15.2	76.0	172742	2	AC023406	AC023406	820	15.2	76.0	199240	10	AC102432	AC102432 Mus muscu
748	15.2	76.0	172962	10	AL713865	Mouse DNA	821	15.2	76.0	199559	2	AL954362	AL954362 Mouse DNA
749	15.2	76.0	173394	9	CNS00001	AL049834	822	15.2	76.0	199923	10	AC122345	AC122345 Mus muscu

C 823	15.2	76.0	199946	9	AC087272	AC087272 Homo sapi	C 896	15.2	76.0	228492	10	AC132058	AC132058 Rattus no
C 824	15.2	76.0	200350	10	AC140208	AC140208 Mus muscu	C 897	15.2	76.0	228647	2	AC134289	AC134289 Rattus no
C 825	15.2	76.0	200785	2	AC108839	AC108839 Mus muscu	C 898	15.2	76.0	228827	2	AC142364	AC142364 Rattus no
C 826	15.2	76.0	200979	9	AC090208	AC090208 Homo sapi	C 899	15.2	76.0	230525	2	AC096982	AC096982 Rattus no
C 827	15.2	76.0	201412	9	AC096921	AC096921 Homo sapi	C 900	15.2	76.0	230734	2	AC115315	AC115315 Rattus no
C 828	15.2	76.0	201500	10	AL645807	AL645807 Mouse DNA	C 901	15.2	76.0	230738	2	AC129135	AC129135 Rattus no
C 829	15.2	76.0	202645	9	AL513330	AL513330 Homo sapi	C 902	15.2	76.0	230892	10	AC131721	AC131721 Rattus no
C 830	15.2	76.0	202661	9	AC143366	AC143366 Pan trogl	C 903	15.2	76.0	230960	2	AC129802	AC129802 Rattus no
C 831	15.2	76.0	202991	10	AC149223	AC149223 Mus muscu	C 904	15.2	76.0	231542	2	AC094709	AC094709 Rattus no
C 832	15.2	76.0	205374	10	AC147160	AC147160 Mus muscu	C 905	15.2	76.0	231671	2	AC111392	AC111392 Rattus no
C 833	15.2	76.0	205621	10	AC149223	AC149223 Mus muscu	C 906	15.2	76.0	232155	2	AC118152	AC118152 Rattus no
C 834	15.2	76.0	205656	2	AC067908	AC067908 Mus muscu	C 907	15.2	76.0	232193	2	AC130163	AC130163 Rattus no
C 835	15.2	76.0	205691	2	AC067908	AC067908 Mus muscu	C 908	15.2	76.0	232473	2	AC098461	AC098461 Rattus no
C 836	15.2	76.0	205737	2	AC087227	AC087227 Mus muscu	C 909	15.2	76.0	233476	2	AC130094	AC130094 Rattus no
C 837	15.2	76.0	205972	2	AC021445	AC021445 Mus muscu	C 910	15.2	76.0	233476	10	AC107662	AC107662 Rattus no
C 838	15.2	76.0	206112	2	AC143415	AC143415 Macaca mu	C 911	15.2	76.0	234778	2	AC097097	AC097097 Rattus no
C 839	15.2	76.0	206310	9	AY371697	AY371697 Homo sapi	C 912	15.2	76.0	234804	2	AC133409	AC133409 Rattus no
C 840	15.2	76.0	206459	2	AC150643	AC150643 Bos tauru	C 913	15.2	76.0	234960	2	AC130922	AC130922 Rattus no
C 841	15.2	76.0	206748	10	AL807804	AL807804 Mouse DNA	C 914	15.2	76.0	235544	10	AC138288	AC138288 Rattus no
C 842	15.2	76.0	206810	2	AC144819	AC144819 Mus muscu	C 915	15.2	76.0	235623	2	AC096613	AC096613 Rattus no
C 843	15.2	76.0	207344	2	AC129235	AC129235 Rattus no	C 916	15.2	76.0	236836	2	AC108530	AC108530 Rattus no
C 844	15.2	76.0	207901	2	AC147043	AC147043 Pan trogl	C 917	15.2	76.0	237859	2	AC124899	AC124899 Rattus no
C 845	15.2	76.0	208293	2	AC127052	AC127052 Rattus no	C 918	15.2	76.0	238061	2	AC109500	AC109500 Rattus no
C 846	15.2	76.0	208311	2	AC105301	AC105301 Mus muscu	C 919	15.2	76.0	238100	2	AC121201	AC121201 Rattus no
C 847	15.2	76.0	208526	2	AC149228	AC149228 Pan trogl	C 920	15.2	76.0	238521	2	AC078932	AC078932 Rattus no
C 848	15.2	76.0	208661	10	AC072048	AC072048 Mus muscu	C 921	15.2	76.0	238658	10	AC079818	AC079818 Rattus no
C 849	15.2	76.0	208757	10	AC100752	AC100752 Mus muscu	C 922	15.2	76.0	238937	2	AC094319	AC094319 Rattus no
C 850	15.2	76.0	209151	2	AC092090	AC092090 Canis fam	C 923	15.2	76.0	239125	2	AC094319	AC094319 Rattus no
C 851	15.2	76.0	209181	2	AC127124	AC127124 Rattus no	C 924	15.2	76.0	239718	2	AC101774	AC101774 Rattus no
C 852	15.2	76.0	209276	2	CR388041	CR388041 Danio rer	C 925	15.2	76.0	240216	2	AC111721	AC111721 Rattus no
C 853	15.2	76.0	211254	2	AC129726	AC129726 Rattus no	C 926	15.2	76.0	241031	2	AC113750	AC113750 Rattus no
C 854	15.2	76.0	211509	9	CNS08CBD	AL928654 Human chr	C 927	15.2	76.0	242907	2	AC106660	AC106660 Rattus no
C 855	15.2	76.0	211546	2	AC135554	AC135554 Gallus ga	C 928	15.2	76.0	243182	2	AC120744	AC120744 Rattus no
C 856	15.2	76.0	211696	2	AC099580	AC099580 Mus muscu	C 929	15.2	76.0	243360	2	AC109684	AC109684 Rattus no
C 857	15.2	76.0	212208	2	AC079024	AC079024 Mus muscu	C 930	15.2	76.0	243768	2	AC125725	AC125725 Rattus no
C 858	15.2	76.0	213181	10	AC084800	AC084800 Mus muscu	C 931	15.2	76.0	244105	2	AC098857	AC098857 Rattus no
C 859	15.2	76.0	213338	2	AC141524	AC141524 Rattus no	C 932	15.2	76.0	244946	2	AC111788	AC111788 Mus muscu
C 860	15.2	76.0	213354	2	AC091030	AC091030 Homo sapi	C 933	15.2	76.0	245136	2	AC111706	AC111706 Rattus no
C 861	15.2	76.0	213826	10	AC131589	AC131589 Mus muscu	C 934	15.2	76.0	245794	2	AC125747	AC125747 Rattus no
C 862	15.2	76.0	214272	2	EX927135	EX927135 Danio rer	C 935	15.2	76.0	245794	2	AC094941	AC094941 Rattus no
C 863	15.2	76.0	214302	2	AC115639	AC115639 Rattus no	C 936	15.2	76.0	245965	2	AC129993	AC129993 Rattus no
C 864	15.2	76.0	214457	10	AC092091	AC092091 Canis fam	C 937	15.2	76.0	245965	2	AC137050	AC137050 Rattus no
C 865	15.2	76.0	214625	2	AC105305	AC105305 Mus muscu	C 938	15.2	76.0	246402	2	AC131535	AC131535 Rattus no
C 866	15.2	76.0	215697	2	AC103050	AC103050 Rattus no	C 939	15.2	76.0	246402	2	AC097034	AC097034 Rattus no
C 867	15.2	76.0	215876	2	AC115909	AC115909 Mus muscu	C 940	15.2	76.0	246635	2	AC094740	AC094740 Rattus no
C 868	15.2	76.0	216078	2	AC115165	AC115165 Rattus no	C 941	15.2	76.0	247306	2	AC131445	AC131445 Rattus no
C 869	15.2	76.0	216252	2	AC135266	AC135266 Rattus no	C 942	15.2	76.0	248567	2	AC128997	AC128997 Rattus no
C 870	15.2	76.0	216736	2	AC124080	AC124080 Mus muscu	C 943	15.2	76.0	250324	2	AC130555	AC130555 Rattus no
C 871	15.2	76.0	216862	2	AC137412	AC137412 Rattus no	C 944	15.2	76.0	251997	2	AC128370	AC128370 Rattus no
C 872	15.2	76.0	216932	10	AC090656	AC090656 Mus muscu	C 945	15.2	76.0	252129	2	AC133975	AC133975 Rattus no
C 873	15.2	76.0	216999	5	AC136741	AC136741 Mus muscu	C 946	15.2	76.0	252561	10	AC102091	AC102091 Mus muscu
C 874	15.2	76.0	217065	2	EX601644	EX601644 zebrafish	C 947	15.2	76.0	253401	2	AC127337	AC127337 Rattus no
C 875	15.2	76.0	217193	2	AC110975	AC110975 Rattus no	C 948	15.2	76.0	253401	1	AL627269	AL627269 Salmo nel
C 876	15.2	76.0	217325	10	AC115360	AC115360 Mus muscu	C 949	15.2	76.0	254050	2	AC109427	AC109427 Rattus no
C 877	15.2	76.0	217801	2	AC134164	AC134164 Rattus no	C 950	15.2	76.0	254388	2	AC109427	AC109427 Rattus no
C 878	15.2	76.0	219121	2	AC134718	AC134718 Rattus no	C 951	15.2	76.0	254926	2	AC130105	AC130105 Rattus no
C 879	15.2	76.0	220678	2	AC023857	AC023857 Homo sapi	C 952	15.2	76.0	256019	9	AC016525	AC016525 Homo sapi
C 880	15.2	76.0	221460	2	AC126434	AC126434 Mus muscu	C 953	15.2	76.0	256320	2	AC132701	AC132701 Rattus no
C 881	15.2	76.0	221516	2	AC095936	AC095936 Rattus no	C 954	15.2	76.0	256643	2	AC122567	AC122567 Rattus no
C 882	15.2	76.0	221621	2	AC113317	AC113317 Mus muscu	C 955	15.2	76.0	259650	2	AC107176	AC107176 Rattus no
C 883	15.2	76.0	221722	2	AC130567	AC130567 Rattus no	C 956	15.2	76.0	259692	2	AC142216	AC142216 Mus muscu
C 884	15.2	76.0	221899	2	AC079548	AC079548 Mus muscu	C 957	15.2	76.0	260909	2	AC105865	AC105865 Rattus no
C 885	15.2	76.0	222498	2	AC094004	AC094004 Rattus no	C 958	15.2	76.0	262048	10	AC093483	AC093483 Mus muscu
C 886	15.2	76.0	222856	2	AC128437	AC128437 Rattus no	C 959	15.2	76.0	262061	2	AC117172	AC117172 Rattus no
C 887	15.2	76.0	223041	10	AC108942	AC108942 Mus muscu	C 960	15.2	76.0	262363	2	AC105703	AC105703 Rattus no
C 888	15.2	76.0	223850	2	AC116397	AC116397 Mus muscu	C 961	15.2	76.0	263151	2	AC128819	AC128819 Rattus no
C 889	15.2	76.0	223859	2	EX927287	EX927287 Danio rer	C 962	15.2	76.0	264033	1	AC133247	AC133247 Rattus no
C 890	15.2	76.0	224808	2	AC094882	AC094882 Rattus no	C 963	15.2	76.0	266050	2	AC090988	AC090988 Sulfolobu
C 891	15.2	76.0	224865	2	AC125933	AC125933 Rattus no	C 964	15.2	76.0	266099	2	AC094795	AC094795 Rattus no
C 892	15.2	76.0	226259	2	AC129784	AC129784 Rattus no	C 965	15.2	76.0	266676	2	AC115671	AC115671 Rattus no
C 893	15.2	76.0	227182	2	AC099283	AC099283 Rattus no	C 966	15.2	76.0	266726	2	AC112860	AC112860 Rattus no
C 894	15.2	76.0	227341	2	AC127771	AC127771 Rattus no	C 967	15.2	76.0	267054	2	AC094354	AC094354 Rattus no
C 895	15.2	76.0	228108	2	AC142063	AC142063 Rattus no	C 968	15.2	76.0	267182	2	AC116720	AC116720 Mus muscu

969 15.2 76.0 267546 2 AC125740 Rattus no
 970 15.2 76.0 267762 2 AC120693 Rattus no
 971 15.2 76.0 269471 2 AC134107 Rattus no
 972 15.2 76.0 270654 2 AC096084 Rattus no
 973 15.2 76.0 270699 2 AC133821 Rattus no
 974 15.2 76.0 271701 2 AC129343 Rattus no
 975 15.2 76.0 272100 3 AE003840 Rattus no
 976 15.2 76.0 272169 2 AC128071 Rattus no
 977 15.2 76.0 272493 2 AC121671 Rattus no
 978 15.2 76.0 278087 2 AC098278 Rattus no
 979 15.2 76.0 281197 2 AC114571 Rattus no
 980 15.2 76.0 283631 2 AC112536 Rattus no
 981 15.2 76.0 284268 2 AE003565 Rattus no
 982 15.2 76.0 285524 2 AC105802 Rattus no
 983 15.2 76.0 287526 2 AC117308 Rattus no
 984 15.2 76.0 290319 1 AE016982 Rattus no
 985 15.2 76.0 295463 2 AC095332 Rattus no
 986 15.2 76.0 295760 2 AC125718 Rattus no
 987 15.2 76.0 296914 2 AC127750 Rattus no
 988 15.2 76.0 297581 2 AC120937 Rattus no
 989 15.2 76.0 298531 2 AC091363 Rattus no
 990 15.2 76.0 298782 2 AC006844 Rattus no
 991 15.2 76.0 300029 1 AE015553 Rattus no
 992 15.2 76.0 300029 1 AE016839 Rattus no
 993 15.2 76.0 300275 1 AE016751 Rattus no
 994 15.2 76.0 300750 1 AP001509 Rattus no
 995 15.2 76.0 309052 3 AE003425 Rattus no
 996 15.2 76.0 312366 2 AC068766 Rattus no
 997 15.2 76.0 322778 2 AC105550 Rattus no
 998 15.2 76.0 326301 1 NM6622491 Rattus no
 999 15.2 76.0 327507 2 AC096293 Rattus no
 1000 15.2 76.0 332760 2 AC147293 Rattus no

ALIGNMENTS

RESULT 1
 AX481360 20 bp DNA linear
 DEFINITION Sequence 7 from Patent EP1225232.
 ACCESSION AX481360
 VERSION AX481360.1 GI:22316281

SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE 1 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 RUBIN, B.Y. and ANDERSON, S.L.
 TITLE Detection of mutations in a gene encoding Ikappab
 JOURNAL Kinase-complex-associated protein to diagnose familial dysautonomia
 Patent: EP 1225232-A 7 24-JUL-2002;
 RUBIN, Berish Y. (US); ANDERSON, Silvia L. (US)
 Location/Qualifiers

FEATURES
 source 1..20
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"

ORIGIN

Query Match 100.0%; Score 20; DB 6; Length 20;
 Best Local Similarity 100.0%; Pred. No. 7.6; 0; Indels 0; Gaps 0;
 Matches 20; Conservative 0; Mismatches 0;

QY 1 AGTCGCAACAGTACATG 20
 Db 1 AGTCGCAACAGTACATG 20

RESULT 2
 AX676048/c 66479 bp DNA linear PAT 27-MAR-2003

DEFINITION Sequence 1 from Patent WO02059381.
 ACCESSION AX676048
 VERSION AX676048.1 GI:29333739
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 Slangenhuysen, S. and Gusejra, J.F.
 TITLE Gene for identifying individuals with familial dysautonomia
 JOURNAL Patent: WO 02059381-A 1 01-AUG-2002;
 The General Hospital Corporation (US)
 Location/Qualifiers

FEATURES

source 1..66479
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"

ORIGIN

Query Match 100.0%; Score 20; DB 6; Length 66479;
 Best Local Similarity 100.0%; Pred. No. 5.2; 0; Indels 0; Gaps 0;
 Matches 20; Conservative 0; Mismatches 0;

QY 1 AGTCGCAACAGTACATG 20
 Db 34222 AGTCGCAACAGTACATG 34203

RESULT 3
 AL359692 78376 bp DNA linear PRI 18-JUL-2001
 LOCUS Human DNA sequence from clone RP11-3J11 on chromosome 9, complete
 DEFINITION sequence.

ACCESSION AL359692
 VERSION AL359692.9 GI:14970800
 KEYWORDS HTG.

SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE 1 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 AUTHORS Sycamore, N.
 TITLE Direct Submission
 JOURNAL Submitted (18-JUL-2001) Sanger Centre, Hinxton, Cambridgeshire,
 CB10 1SA, UK. E-mail enquiries: humgeny@sanger.ac.uk
 requests: clonerequests@sanger.ac.uk

COMMENT

On Jul 19, 2001 this sequence version replaced gi:14715342.
 During sequence assembly data is compared from overlapping clones.
 Where differences are found these are annotated as variations
 together with a note of the overlapping clone name. Note that the
 variation annotation may not be found in the sequence submission
 corresponding to the overlapping clone, as we submit sequences with
 only a small overlap as described above.
 This sequence was finished as follows unless otherwise noted: all
 regions were either double-stranded or sequenced with an alternate
 chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such
 as compressions and repeats; all regions were covered by at least
 one plasmid subclone or more than one M13 subclone; and the
 assembly was confirmed by restriction digest. The following
 abbreviations are used to associate primary accession numbers given
 in the feature table with their source databases: Emi, EMBL; Swi,
 SWISSPROT; Tr, TREMBL; Wp, WORMPEP; Information on the WORMPEP
 database can be found at
 http://www.sanger.ac.uk/Projects/C_elegans/wormep This sequence
 was generated from part of bacterial clone contigs of human
 chromosome 9, constructed by the Sanger Centre Chromosome 9 Mapping
 Group. Further information can be found at
 http://www.sanger.ac.uk/HGP/Chr9
 RP11-3J11 is from the library RP11-11.1 constructed by the group of
 Pieter de Jong. For further details see
 http://www.chori.org/bacpac/home.htm

VECTOR: pBACE3.6
IMPORTANT: This sequence is not the entire insert of clone RP11-3J11. It may be shorter because we sequence overlapping sections only once, except for a 100 base overlap. The true left end of clone RP11-3J11 is at 1 in this sequence. The true left end of clone RP11-115022 is at 76377 in this sequence. The true right end of clone RP11-3J11 is at 76382 in this sequence.

FEATURES Location/Qualifiers

```
source
1..78376
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="9"
/clone="RP11-3J11"
/clone_1fb="RP11-11.1"
7..829
/feature="MIR repeat: matches 5343..6155 of consensus"
repeat_region
991..13328
/feature="MER46B repeat: matches 3..231 of consensus"
repeat_region
2199..2234
/feature="18 copies 2 mer tt 86% conserved"
repeat_region
3191..3349
/feature="AluSg/x repeat: matches 153..311 of consensus"
repeat_region
4250..4537
/feature="AluSx repeat: matches 1..288 of consensus"
repeat_region
5314..5579
/feature="AluYb repeat: matches 1..299 of consensus"
repeat_region
5693..5863
/feature="L1MC4 repeat: matches 6481..6679 of consensus"
repeat_region
5881..6013
/feature="AluYb repeat: matches 1..135 of consensus"
repeat_region
6533..6644
/feature="MIR repeat: matches 92..211 of consensus"
repeat_region
7645..7799
/feature="MIR repeat: matches 3..363 of consensus"
repeat_region
9165..9774
/feature="MER21B repeat: matches 179..787 of consensus"
repeat_region
9775..9848
/feature="MER84 repeat: matches 1..76 of consensus"
repeat_region
9849..10232
/feature="MER84 repeat: matches 3..373 of consensus"
repeat_region
10233..10535
/feature="AluSx repeat: matches 1..302 of consensus"
repeat_region
10536..10668
/feature="MER84 repeat: matches 373..508 of consensus"
repeat_region
10669..10846
/feature="MER21B repeat: matches 7..185 of consensus"
repeat_region
11168..11901
/feature="L1PA14 repeat: matches 5403..6149 of consensus"
repeat_region
12172..12468
/feature="AluSg repeat: matches 1..297 of consensus"
repeat_region
13074..13362
/feature="AluSx repeat: matches 3..294 of consensus"
repeat_region
13411..13533
/feature="MIR repeat: matches 117..239 of consensus"
repeat_region
13616..13669
/feature="L2 repeat: matches 2445..2498 of consensus"
repeat_region
13733..13841
/feature="MER53 repeat: matches 89..188 of consensus"
repeat_region
13996..14113
/feature="AluSg/x repeat: matches 178..295 of consensus"
repeat_region
14769..15142
/feature="MSTRA repeat: matches 1..423 of consensus"
repeat_region
15274..15333
/feature="MER5A repeat: matches 14..74 of consensus"
repeat_region
15480..15600
/feature="MER5B repeat: matches 54..174 of consensus"
repeat_region
15640..15689
/feature="MSTB repeat: matches 1..49 of consensus"
repeat_region
15881..16180
/feature="AluYb repeat: matches 1..302 of consensus"
repeat_region
17189..17296
```

```
repeat_region
17614..17738
/feature="MER5B repeat: matches 54..174 of consensus"
repeat_region
17769..18234
/feature="L1MD1 repeat: matches 5748..6223 of consensus"
repeat_region
18235..18529
/feature="AluYb repeat: matches 1..292 of consensus"
repeat_region
18530..19071
/feature="L1MD1 repeat: matches 5238..5748 of consensus"
repeat_region
19132..19432
/feature="AluSg repeat: matches 1..313 of consensus"
repeat_region
19441..19472
/feature="16 copies 2 mer tt 84% conserved"
repeat_region
19498..19859
/feature="THB1A repeat: matches 1..353 of consensus"
repeat_region
19860..21021
/feature="THB1A-internal repeat: matches 427..1580 of consensus"
repeat_region
21022..21314
/feature="AluSg repeat: matches 1..302 of consensus"
repeat_region
21315..21711
/feature="THB1A-internal repeat: matches 29..427 of consensus"
repeat_region
21718..21818
/feature="THB1B repeat: matches 265..364 of consensus"
repeat_region
21813..21845
/feature="Alu repeat: matches 3..35 of consensus"
repeat_region
21878..22184
/feature="AluSg repeat: matches 1..311 of consensus"
repeat_region
22578..22715
/feature="MIR repeat: matches 74..210 of consensus"
repeat_region
22693..22848
/feature="L2 repeat: matches 2572..2745 of consensus"
repeat_region
24368..24434
/feature="MIR repeat: matches 71..137 of consensus"
repeat_region
25259..25338
/feature="L2 repeat: matches 2605..2695 of consensus"
repeat_region
25474..25775
/feature="AluSg repeat: matches 1..301 of consensus"
repeat_region
26871..27694
/feature="CpG island"
/feature="evidence=not_experimental"
repeat_region
28479..28806
/feature="L2 repeat: matches 2157..2496 of consensus"
repeat_region
28807..29114
/feature="AluSg repeat: matches 1..305 of consensus"
repeat_region
29115..29296
/feature="L2 repeat: matches 1963..2157 of consensus"
repeat_region
29748..30050
/feature="AluSg repeat: matches 1..303 of consensus"
repeat_region
30475..30548
/feature="MIR repeat: matches 33..107 of consensus"
repeat_region
30654..30701
/feature="24 copies 2 mer ac 95% conserved"
repeat_region
31870..32343
/feature="L1MD1 repeat: matches 16..505 of consensus"
repeat_region
33199..33414
/feature="L2 repeat: matches 2116..2353 of consensus"
repeat_region
33415..33568
/feature="MER58B repeat: matches 186..340 of consensus"
repeat_region
33658..33918
/feature="L2 repeat: matches 2403..2710 of consensus"
repeat_region
37241..37362
/feature="Char11e4 repeat: matches 1823..1956 of consensus"
repeat_region
37669..37895
/feature="MER20 repeat: matches 1..218 of consensus"
repeat_region
38589..38875
/feature="AluSg repeat: matches 1..310 of consensus"
repeat_region
39767..39947
/feature="MSTB repeat: matches 1..183 of consensus"
repeat_region
40006..40255
/feature="match: STS: Em:G14836"
misc_feature
40477..40798
```

```

repeat_region /note="match: STS: Em:G21606"
41261..41395
/note="LINE2 repeat: matches 6041..6161 of consensus"
repeat_region 41739..41805
/note="TIGER2 repeat: matches 2653..2718 of consensus"
repeat_region 41806..42093
/note="TIGER2 repeat: matches 1..300 of consensus"
repeat_region 42094..43429
/note="TIGER2 repeat: matches 1302..2653 of consensus"
repeat_region 43429..44232
/note="TIGER2 repeat: matches 297..1113 of consensus"
repeat_region 44353..44664

```

```

Query Match 100.0%; Score 20; DB 9; Length 78376;
Best Local Similarity 100.0%; Pred. No. 5.1;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

RESULT 4

AC148981_2/c

WPCOMMENT

Sequence split into 4 fragments

Fragment Name

Begin

End

AC148981_0 1 110000

AC148981_1 100001 210000

AC148981_2 200001 310000

AC148981_3 300001 359936

Continuation (3 of 4) of AC148981 from base 200001 (AC148981 Mus musculus chromosome 7 c

```

Query Match 87.0%; Score 17.4; DB 2; Length 110000;
Best Local Similarity 94.7%; Pred. No. 1.5e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

QY 1 AGTGGCAACAGTACATG 19

Db 92061 AGTCCCAACAGTACATG 92043

```

RESULT 5
AC073603/c 135276 bp DNA linear HTG 06-SEP-2000
LOCUS Mus musculus clone RP23-6814, WORKING DRAFT SEQUENCE, 13 unordered
pieces.
AC073603
AC073603.3 GI:9972336
AC073603.3 GI:9972336
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus

```

```

REFERENCE
AUTHORS McCombie, W.R., Baker, J.P., Bahret, A., Bal, H., Debia, N.N., de la
Bastide, M., Huang, B.N., King, L., Kirchoff, K.A., Miller, B.,
Nascimento, L.U., O'Shaughnessy, A.L., Preston, R.R., Rodriguez, M.A.,
Shah, R.S., Shekher, M., Spiegel, L.A., Toth, K. and Val, M.D.
TITLE Mouse Genomic Sequence
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 135276)
AUTHORS McCombie, W.R.
JOURNAL Direct Submission

```

```

COMMENT
Submitted (26-JUN-2000) Lita Annenberg Hazen Genome Sequencing
Center, Cold Spring Harbor Laboratory, 1 Bungtown Road, Cold Spring
Harbor, NY 11724, USA
On Sep 6, 2000 this sequence version replaced gi:9797698.
----- Genome Center
Center: Lita Annenberg Hazen Genome Center, Cold Spring Harbor
Laboratory
Center code: CSHL
Web site: http://www.cshl.org/genseq

```

```

Contact: mcombie@cshl.org
----- Project Information
Center project name: RP23-6814
Center clone name: RP23-6814
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 13 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

```

```

1 19589: contig of 19589 bp in length
* 19590 20135: gap of unknown length
* 20140 39580: contig of 19441 bp in length
* 39581 40130: gap of unknown length
* 40131 58318: contig of 18188 bp in length
* 58319 58868: gap of unknown length
* 58869 75533: contig of 16665 bp in length
* 75534 76083: gap of unknown length
* 76084 92119: contig of 16036 bp in length
* 92120 92669: gap of unknown length
* 92670 101790: contig of 9121 bp in length
* 101791 102340: gap of unknown length
* 102341 109258: contig of 6918 bp in length
* 109259 109808: gap of unknown length
* 109809 116229: contig of 6421 bp in length
* 116230 116779: gap of unknown length
* 116780 123063: contig of 6284 bp in length
* 123064 123612: gap of unknown length
* 123613 127541: contig of 3929 bp in length
* 127542 128090: gap of unknown length
* 128091 131354: contig of 3264 bp in length
* 131355 131903: gap of unknown length
* 131904 134618: contig of 2715 bp in length
* 134619 135167: gap of unknown length
* 135168 135276: contig of 109 bp in length.
Location/Qualifiers
1..135276
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/clone="RP23-6814"

```

FEATURES

source

```

ORIGIN
Query Match 87.0%; Score 17.4; DB 2; Length 135276;
Best Local Similarity 94.7%; Pred. No. 1.4e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

QY 1 AGTGGCAACAGTACATG 19

Db 2982 AGTCCCAACAGTACATG 2964

```

RESULT 6
AC140732/c 151366 bp DNA linear HTG 27-MAR-2003
LOCUS Rattus norvegicus clone CH230-24B6, WORKING DRAFT SEQUENCE, 17
unordered pieces.
AC140732
AC140732.1 GI:28626532
AC140732.1 GI:28626532
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus

```

```

REFERENCE
AUTHORS Muzny, D., Marie, M., Merzker, M., Lee, A., Abramson, S., Adams, C., Alder, J.,
Allen, C., Allen, H., Aisbrough, S., Amin, A., Arguiano, D.,
Ayalebech, V., Ayagi, A., Ayodeji, M., Baca, E., Baden, H.,
Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F.,

```

Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M., Bryant, N., Buhay, C., Burch, P., Butrell, K., Calderon, E., Cardenas, V., Carter, K., Cavazos, I., Caesar, H., Center, A., Chacco, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J., Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L., Davila, M.L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D., Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Diya, K., Diaper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Evans, K., Egan, A., Escotto, M., Eugene, C., Evans, C.A., Falls, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Fraser, C.M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M., Gebregiorgis, B., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W., Gunaratne, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K., Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, C., Hernandez, R., Hines, S., Hladun, S.L., Hodgson, A., Hogue, M., Hollins, B., Howells, S., Huliy, S., Hume, J., Idlebird, D., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A., Karpathy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kowis, C., Kraft, C.L., Lebow, H., Levant, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lorensunewa, L., Louised, H., Lozano, R.J., Lu, X., Ma, U., Maheshwari, M., Mahindaratne, N., Mahmoud, M., Malloy, K., Mangum, A., Mangum, B., Mapa, P., Martin, K., Martin, R., Martinez, E., Mawhinney, S., McLeod, M., McNeill, T., Meenen, E., Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Munday, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwaokeleneh, O., Okunolu, G., Olarinmwa, A., Pal, S., Parker, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C., Plopper, F., Polidexter, A., Popovic, D., Primus, E., Pu, L., L., Puzo, M., Quiroz, J., Sachin, E., Reeves, K., Regier, M.A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rivers, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S., Sanders, W., Savary, G., Scherer, S., Scott, G., Shatman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C.D., Sma, D., Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J., Steimle, M., Strong, R., Sutton, A., Syarik, A., Taber, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K., Valas, R., Vera, V., Villalana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, D., Warren, R., Wei, X., White, P., Williams, G., Wilson, R., Wlezyk, R., Wooden, R., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, U., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Niederhausern, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O., Weinstock, G. and Gibbs, R.A.

Unpublished
2 (bases 1 to 151366)
Worley, K.C.
Direct Submission
Submitted (01-MAR-2003) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 151366)
Worley, K.C.
Direct Submission
Submitted (27-MAR-2003) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: KEBF
Center clone name: CH230-24B6
----- Summary Statistics
Sequencing vector: Plasmid
Chemistry: Dye-terminator Big Dye 100% of reads
Assembly program: Phred; version 0.990329
Consensus quality: 143638 bases at least Q40
Consensus quality: 145076 bases at least Q30
Consensus quality: 145959 bases at least Q20

Estimated insert size: 143140; sum-of-contigs estimation
Quality coverage: 4x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
* (see <http://www.hgsc.bcm.tmc.edu/docs/genbankdraftdata.html>).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 17 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1	1151: contig of 1151 bp in length
*	1152: gap of unknown length
*	1152: contig of 1075 bp in length
*	2327: gap of unknown length
*	2426: contig of 1161 bp in length
*	2427: gap of unknown length
*	3588: contig of 1809 bp in length
*	3688: gap of unknown length
*	5497: gap of unknown length
*	5597: gap of 3031 bp in length
*	8628: gap of unknown length
*	8727: gap of 4297 bp in length
*	13025: gap of unknown length
*	13125: gap of 7562 bp in length
*	20687: gap of unknown length
*	20786: gap of 7922 bp in length
*	28708: gap of unknown length
*	28709: gap of 8210 bp in length
*	28809: gap of unknown length
*	37018: gap of 8055 bp in length
*	37119: gap of unknown length
*	45173: gap of unknown length
*	45273: gap of 11243 bp in length
*	45274: gap of unknown length
*	55517: gap of unknown length
*	56167: gap of 12817 bp in length
*	69433: gap of unknown length
*	69534: gap of unknown length
*	69537: gap of 10024 bp in length
*	79558: gap of unknown length
*	79558: gap of 17087 bp in length
*	96745: gap of unknown length
*	96844: gap of 13707 bp in length
*	110551: gap of unknown length
*	110552: gap of 17448 bp in length
*	128099: gap of unknown length
*	128100: gap of unknown length
*	151366: contig of 23167 bp in length.

***** Location/Qualifiers *****
1.151366
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-24B6"

ORIGIN
Query Match 87.0% Score 17.4; DB 2; Length 151366;
Best Local Similarity 94.7%; Pred. No. 1.4e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GTGGCAACAGTACATGG 20
DB 103450 GTGGCAACAGTACATGG 103432

RESULT 7
AC099737 202183 bp DNA linear HTG 03-MAR-2004
LOCUS Mus musculus chromosome 17 clone RP23-247M16 map 17, *** SEQUENCING
DEFINITION IN PROGRESSES ***; 8 unordered pieces.
AC099737
AC099737.3 GI:44890998
VERSION HTG; HTGS_PHASE1; HTGS_FULFILL; HTGS_ACTIVEFIN.
KEYWORDS HTG; HTGS (house mouse)
SOURCE
ORGANISM Mus musculus

REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 202183)
Birren, B., Nusbaum, C. and Lander, E.
Mus musculus chromosome 17, clone RP23-247M16
Unpublished
2 (bases 1 to 202183)

Birren, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S., Barna, N., Bastien, V., Boguslavsky, L., Boukhgalter, B., Brown, A., Camarata, J., Campopiano, A., Chang, J., Chazaro, B., Choepel, Y., Colangelo, M., Collins, S., Collamore, A., Cook, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J.S., Dodge, S., Faro, S., Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J., Gardyna, S., Grande, S., Grotz, S., Goyette, M., Graham, L., Grand-Pierre, N., Jones, C., Heatford, A., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jorgensen, C., Kamat, A., Karatas, A., Kellis, C., Larocque, K., Lamazeres, R., Landers, T., Lehoczy, J., Levine, R., Liu, G., Maclean, C., Macdonald, P., Major, J., Margulis, N., Matthews, C., McCarthy, M., McGowan, P., McKernan, K., McPheters, R., Meldrum, J., Menues, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V., Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schnupack, R., Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, M.J., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

TITLE
JOURNAL
REFERENCE
AUTHORS

Submitted (19-NOV-2001) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 202183)

Birren, B., Nusbaum, C., Lander, E., Abouelleil, A., Allen, N., Anderson, M., Arachchi, H.W., Barna, N., Bastien, V., Bloom, T., Boguslavsky, L., Boukhgalter, B., Camarata, J., Chang, J., Choepel, Y., Collamore, A., Cooke, P., Corum, B., Dearellano, K., Diaz, J.S., Dodge, S., Dooley, K., Dorris, L., Erickson, J., Faro, S., Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J., Gardyna, S., Graham, L., Grand-Pierre, N., Hafez, N., Hapcihan, D., Hagos, B., Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kellis, C., Landers, T., Levine, R., Lindblad-Toh, K., Liu, X., Lui, A., Mabbitt, R., Maclean, C., MacDonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M., Meldrum, J., Menues, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Rachupka, A., Ramasamy, U., Raymond, C., Retta, R., Rise, C., Rogov, P., Roman, J., Schauer, S., Schnupack, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H., Venkataraman, V.S., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.
Direct Submission
Submitted (03-MAR-2004) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
On Mar 3, 2004 this sequence version replaced g1:28475922.
All repeats were identified using RepeatMasker:
http://ftp.genome.washington.edu/RW/RepeatMasker.html

Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
Project Information
Center project name: L17629
Center clone name: 247_M_16

* NOTE: This is a 'working draft' sequence. It currently
* consists of 8 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as

* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 39687: contig of 39687 bp in length
* 39688 39787: gap of 100 bp
* 39788 80800: contig of 41013 bp in length
* 80801 80900: gap of 100 bp
* 80901 131062: contig of 50162 bp in length
* 131063 131162: gap of 100 bp
* 131163 140256: contig of 9094 bp in length
* 140257 140356: gap of 100 bp
* 140357 146397: contig of 6641 bp in length
* 146398 147097: gap of 100 bp
* 147098 172292: contig of 25195 bp in length
* 172293 172392: gap of 100 bp
* 172393 191398: contig of 19006 bp in length
* 191399 191498: gap of 100 bp
* 191499 202183: contig of 10685 bp in length.
Location/Qualifiers
1..202183
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="17"
/map="17"
/clone="RP23-247M16"
/clone_1fb="RP23-247M16"
/clone_1fb="RP23-247M16" Female Mouse BAC"

ORIGIN

Query Match 87.0%; Score 17.4; DB 2; Length 202183;
Best Local Similarity 94.7%; Pred. No. 1.4e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AGTCGCAACGCTCAATG 19
Db 128238 AGTCGCAACGCTCAATG 128256

RESULT 8
AC119471/c

LOCUS Rattus norvegicus clone CH230-414D17, WORKING DRAFT SEQUENCE, 5
DEFINITION Rattus norvegicus clone CH230-414D17, WORKING DRAFT SEQUENCE, 5
AC119471
AC119471.5 GI:25098850
VERSION HTG; HTGS PHASE1; HTGS DRAFT; HTGS_FULLTOP.
KEYWORDS Rattus norvegicus (Norway rat)
SOURCE Rattus norvegicus
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
1 (bases 1 to 205681)

Muzny, D., Marle, Metzker, M., Lee, S., Abramson, S., Adams, C., Alder, J., Allen, C., Allen, H., Albrooke, S., Amin, A., Anguiano, D., Anyalbechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H., Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F., Biewald, K., Blair, J., Blankenburg, K., Blyth, P., Brown, W., Bryant, N., Buhay, C., Burck, P., Burrelli, K., Calderon, E., Cardenas, V., Carter, K., Cavazos, I., Cesari, H., Center, A., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J., Cleveland, C., Cockrell, R., Cox, C., Coyote, M., Cree, A., D'Souza, L., Davila, M.L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D., Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Diya, K., Diaper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K., Egan, A., Escotto, M., Eugene, C., Evans, C.A., Falls, T., Fan, G., Fernandez, S., Finley, M., Flanagan, N., Forbes, J., Foster, M., Foster, P., Fraser, C.M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M., Gebregiorgis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W., Gunaratne, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K., Harvey, Y., Havlak, P., Hayes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S.L., Hodgson, A., Hogues, M., Hollins, B., Howells, S., Hulik, S., Hume, J., Idlebird, D., Jackson, A.,

Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A.,
 Karpachy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C.,
 Kowis, C., Kraft, C.L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J.,
 Liu, P., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J.,
 Lorenz, R., Loulsges, H., Lozada, R.J., Lu, X., Ma, J.,
 Maheshwari, M., Mahindratne, M., Mahmood, M., Malloy, K., Mangum, A.,
 Mangun, B., Mapua, P., Martin, K., Martin, R., Martinez, E.,
 Mathew, S., McLeod, M.P., McNeill, T.Z., Meenan, E.,
 Mawlawi, J., McLeod, M.P., Miner, G., Minja, E., Montemayor, J., Moore, S.,
 Morgan, M., Morris, K., Morris, S., Mundasa, M., Murphy, M., Nair, L.,
 Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S.,
 Nwokwelen, O., Okunnu, G., Olarinmose, A., Pal, S., Parks, K.,
 Patel, S., Paul, H., Perez, A., Perez, L., Pfamkoch, C.,
 Pioppert, F., Polidexter, A., Popovic, D., Primus, E., Pu, L.,
 Piazzi, M., Quirce, J., Rachlin, E., Reeves, K., Regier, M.A., Reich, R.,
 Rial, B., Rial, M., Ren, Y., Reuter, M., Richards, S., Riggs, F.,
 Rivers, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S.J.,
 Sanders, W., Savary, G., Scherer, S., Scott, G., Shatsman, S., Shen, H.,
 Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C.D., Smajls, D.,
 Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J.,
 Steimle, M., Strong, R., Sutton, A., Svatek, A., Taber, P., Taylor, C.,
 Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usman, K.,
 Valas, R., Vera, V., Villaseca, D., Waldron, L., Walker, B., Wang, J.,
 Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, P.,
 Williams, G., Willson, R., Wleczky, R., Wooden, H., Worley, K.,
 Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V.,
 Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von
 Niederhausen, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O.,
 Weinstein, G., and Gibbs, R.A.
 Direct Submission
 Unpublished
 2 (bases 1 to 205681)
 Worley, K.C.
 Direct Submission
 Submitted (27-Apr-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 3 (bases 1 to 205681)
 Rat Genome Sequencing Consortium.
 Direct Submission
 Submitted (19-Nov-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 On Nov 19, 2002 this sequence version replaced gi:2390845.
 The sequence in this assembly is a combination of BAC based reads
 and whole genome shotgun sequencing reads assembled using Atlas
 (<http://www.hgsc.bcm.tmc.edu/projects/rac/>). Each contig described
 in the feature table below represents a scaffold in the Atlas
 assembly (a 'contig-scaffold'). Within each contig-scaffold,
 individual sequence contigs are ordered and oriented, and separated
 by sized gaps filled with Ns to the estimated size. The sequence
 may extend beyond the ends of the clone and there may be sequence
 contigs within a contig-scaffold that consist entirely of whole
 genome shotgun sequence reads. Both end sequences and whole genome
 shotgun sequence only contigs will be indicated in the feature
 table.

----- Genome Center
 Center: Baylor College of Medicine
 Center code: BCM
 Web site: <http://www.hgsc.bcm.tmc.edu/>
 Contact: hgsc-help@bcm.tmc.edu
 ----- Project Information
 Center project name: GVEP
 Center clone name: CH230-414D17
 ----- Summary Statistics
 Assembly program: Phrap, version 0.990329
 Consensus quality: 192343 bases at least Q40
 Consensus quality: 19463 bases at least Q30
 Consensus quality: 195894 bases at least Q20
 Estimated insert size: 198896; sum-of-contigs estimation
 Quality coverage: 7x in Q20 bases; sum-of-contigs estimation

 * NOTE: Estimated insert size may differ from sequence length

(see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 5 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

1 196852: contig of 196852 bp in length
 * 196853 196952: gap of unknown length
 * 196953 198215: contig of 1263 bp in length
 * 198216 198315: gap of unknown length
 * 198316 200209: contig of 1894 bp in length
 * 200210 200309: gap of unknown length
 * 200310 202037: contig of 1728 bp in length
 * 202038 202137: gap of unknown length
 * 202138 205681: contig of 3544 bp in length.

FEATURES
 source
 1. 205681
 /organism="Rattus norvegicus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10116"
 /clone="CH230-414D17"
 1. 2063
 /note="wgs_end_extension
 clone_end:77"
 /note="clone_boundary
 clone_end:77"
 site:
 end_sequence:BZ170605"
 complement(195447..196272)
 /note="clone_boundary
 clone_end:5p6
 site:
 end_sequence:BZ170609"

misc_feature
 7809..8647
 /note="clone_boundary
 clone_end:77"
 site:
 end_sequence:BZ170605"
 complement(195447..196272)
 /note="clone_boundary
 clone_end:5p6
 site:
 end_sequence:BZ170609"

misc_feature
 1. 2063
 /note="wgs_end_extension
 clone_end:77"
 /note="clone_boundary
 clone_end:77"
 site:
 end_sequence:BZ170605"
 complement(195447..196272)
 /note="clone_boundary
 clone_end:5p6
 site:
 end_sequence:BZ170609"

misc_feature
 1. 2063
 /note="wgs_end_extension
 clone_end:77"
 /note="clone_boundary
 clone_end:77"
 site:
 end_sequence:BZ170605"
 complement(195447..196272)
 /note="clone_boundary
 clone_end:5p6
 site:
 end_sequence:BZ170609"

misc_feature
 1. 2063
 /note="wgs_end_extension
 clone_end:77"
 /note="clone_boundary
 clone_end:77"
 site:
 end_sequence:BZ170605"
 complement(195447..196272)
 /note="clone_boundary
 clone_end:5p6
 site:
 end_sequence:BZ170609"

misc_feature
 1. 2063
 /note="wgs_end_extension
 clone_end:77"
 /note="clone_boundary
 clone_end:77"
 site:
 end_sequence:BZ170605"
 complement(195447..196272)
 /note="clone_boundary
 clone_end:5p6
 site:
 end_sequence:BZ170609"

misc_feature
 1. 2063
 /note="wgs_end_extension
 clone_end:77"
 /note="clone_boundary
 clone_end:77"
 site:
 end_sequence:BZ170605"
 complement(195447..196272)
 /note="clone_boundary
 clone_end:5p6
 site:
 end_sequence:BZ170609"

misc_feature
 1. 2063
 /note="wgs_end_extension
 clone_end:77"
 /note="clone_boundary
 clone_end:77"
 site:
 end_sequence:BZ170605"
 complement(195447..196272)
 /note="clone_boundary
 clone_end:5p6
 site:
 end_sequence:BZ170609"

misc_feature
 1. 2063
 /note="wgs_end_extension
 clone_end:77"
 /note="clone_boundary
 clone_end:77"
 site:
 end_sequence:BZ170605"
 complement(195447..196272)
 /note="clone_boundary
 clone_end:5p6
 site:
 end_sequence:BZ170609"

misc_feature
 1. 2063
 /note="wgs_end_extension
 clone_end:77"
 /note="clone_boundary
 clone_end:77"
 site:
 end_sequence:BZ170605"
 complement(195447..196272)
 /note="clone_boundary
 clone_end:5p6
 site:
 end_sequence:BZ170609"

misc_feature
 1. 2063
 /note="wgs_end_extension
 clone_end:77"
 /note="clone_boundary
 clone_end:77"
 site:
 end_sequence:BZ170605"
 complement(195447..196272)
 /note="clone_boundary
 clone_end:5p6
 site:
 end_sequence:BZ170609"

misc_feature
 1. 2063
 /note="wgs_end_extension
 clone_end:77"
 /note="clone_boundary
 clone_end:77"
 site:
 end_sequence:BZ170605"
 complement(195447..196272)
 /note="clone_boundary
 clone_end:5p6
 site:
 end_sequence:BZ170609"

misc_feature
 1. 2063
 /note="wgs_end_extension
 clone_end:77"
 /note="clone_boundary
 clone_end:77"
 site:
 end_sequence:BZ170605"
 complement(195447..196272)
 /note="clone_boundary
 clone_end:5p6
 site:
 end_sequence:BZ170609"

misc_feature
 1. 2063
 /note="wgs_end_extension
 clone_end:77"
 /note="clone_boundary
 clone_end:77"
 site:
 end_sequence:BZ170605"
 complement(195447..196272)
 /note="clone_boundary
 clone_end:5p6
 site:
 end_sequence:BZ170609"

misc_feature
 1. 2063
 /note="wgs_end_extension
 clone_end:77"
 /note="clone_boundary
 clone_end:77"
 site:
 end_sequence:BZ170605"
 complement(195447..196272)
 /note="clone_boundary
 clone_end:5p6
 site:
 end_sequence:BZ170609"

misc_feature
 1. 2063
 /note="wgs_end_extension
 clone_end:77"
 /note="clone_boundary
 clone_end:77"
 site:
 end_sequence:BZ170605"
 complement(195447..196272)
 /note="clone_boundary
 clone_end:5p6
 site:
 end_sequence:BZ170609"

misc_feature
 1. 2063
 /note="wgs_end_extension
 clone_end:77"
 /note="clone_boundary
 clone_end:77"
 site:
 end_sequence:BZ170605"
 complement(195447..196272)
 /note="clone_boundary
 clone_end:5p6
 site:
 end_sequence:BZ170609"

misc_feature
 1. 2063
 /note="wgs_end_extension
 clone_end:77"
 /note="clone_boundary
 clone_end:77"
 site:
 end_sequence:BZ170605"
 complement(195447..196272)
 /note="clone_boundary
 clone_end:5p6
 site:
 end_sequence:BZ170609"

misc_feature
 1. 2063
 /note="wgs_end_extension
 clone_end:77"
 /note="clone_boundary
 clone_end:77"
 site:
 end_sequence:BZ170605"
 complement(195447..196272)
 /note="clone_boundary
 clone_end:5p6
 site:
 end_sequence:BZ170609"

misc_feature
 1. 2063
 /note="wgs_end_extension
 clone_end:77"
 /note="clone_boundary
 clone_end:77"
 site:
 end_sequence:BZ170605"
 complement(195447..196272)
 /note="clone_boundary
 clone_end:5p6
 site:
 end_sequence:BZ170609"

misc_feature
 1. 2063
 /note="wgs_end_extension
 clone_end:77"
 /note="clone_boundary
 clone_end:77"
 site:
 end_sequence:BZ170605"
 complement(195447..196272)
 /note="clone_boundary
 clone_end:5p6
 site:
 end_sequence:BZ170609"

misc_feature
 1. 2063
 /note="wgs_end_extension
 clone_end:77"
 /note="clone_boundary
 clone_end:77"
 site:
 end_sequence:BZ170605"
 complement(195447..196272)
 /note="clone_boundary
 clone_end:5p6
 site:
 end_sequence:BZ170609"

misc_feature
 1. 2063
 /note="wgs_end_extension
 clone_end:77"
 /note="clone_boundary
 clone_end:77"
 site:
 end_sequence:BZ170605"
 complement(195447..196272)
 /note="clone_boundary
 clone_end:5p6
 site:
 end_sequence:BZ170609"

misc_feature
 1. 2063
 /note="wgs_end_extension
 clone_end:77"
 /note="clone_boundary
 clone_end:77"
 site:
 end_sequence:BZ170605"
 complement(195447..196272)
 /note="clone_boundary
 clone_end:5p6
 site:
 end_sequence:BZ170609"

misc_feature
 1. 2063
 /note="wgs_end_extension
 clone_end:77"
 /note="clone_boundary
 clone_end:77"
 site:
 end_sequence:BZ170605"
 complement(195447..196272)
 /note="clone_boundary
 clone_end:5p6
 site:
 end_sequence:BZ170609"

misc_feature
 1. 2063
 /note="wgs_end_extension
 clone_end:77"
 /note="clone_boundary
 clone_end:77"
 site:
 end_sequence:BZ170605"
 complement(195447..196272)
 /note="clone_boundary
 clone_end:5p6
 site:
 end_sequence:BZ170609"

misc_feature
 1. 2063
 /note="wgs_end_extension
 clone_end:77"
 /note="clone_boundary
 clone_end:77"
 site:
 end_sequence:BZ170605"
 complement(195447..196272)
 /note="clone_boundary
 clone_end:5p6
 site:
 end_sequence:BZ170609"

misc_feature
 1. 2063
 /note="wgs_end_extension
 clone_end:77"
 /note="clone_boundary
 clone_end:77"
 site:
 end_sequence:BZ170605"
 complement(195447..196272)
 /note="clone_boundary
 clone_end:5p6
 site:
 end_sequence:BZ170609"

misc_feature
 1. 2063
 /note="wgs_end_extension
 clone_end:77"
 /note="clone_boundary
 clone_end:77"
 site:
 end_sequence:BZ170605"
 complement(195447..196272)
 /note="clone_boundary
 clone_end:5p6
 site:
 end_sequence:BZ170609"

misc_feature
 1. 2063
 /note="wgs_end_extension
 clone_end:77"
 /note="clone_boundary
 clone_end:77"
 site:
 end_sequence:BZ170605"
 complement(195447..196272)
 /note="clone_boundary
 clone_end:5p6
 site:
 end_sequence:BZ170609"

misc_feature
 1. 2063
 /note="wgs_end_extension
 clone_end:77"
 /note="clone_boundary
 clone_end:77"
 site:
 end_sequence:BZ170605"
 complement(195447..196272)
 /note="clone_boundary
 clone_end:5p6
 site:
 end_sequence:BZ170609"

misc_feature
 1. 2063
 /note="wgs_end_extension
 clone_end:77"
 /note="clone_boundary
 clone_end:77"
 site:
 end_sequence:BZ170605"
 complement(195447..196272)
 /note="clone_boundary
 clone_end:5p6
 site:
 end_sequence:BZ170609"

misc_feature
 1. 2063
 /note="wgs_end_extension
 clone_end:77"
 /note="clone_boundary
 clone_end:77"
 site:
 end_sequence:BZ170605"
 complement(195447..196272)
 /note="clone_boundary
 clone_end:5p6
 site:
 end_sequence:BZ170609"

misc_feature
 1. 2063
 /note="wgs_end_extension
 clone_end:77"
 /note="clone_boundary
 clone_end:77"
 site:
 end_sequence:BZ170605"
 complement(195447..196272)
 /note="clone_boundary
 clone_end:5p6
 site:
 end_sequence:BZ170609"

misc_feature
 1. 2063
 /note="wgs_end_extension
 clone_end:77"
 /note="clone_boundary
 clone_end:77"
 site:
 end_sequence:BZ170605"
 complement(195447..196272)
 /note="clone_boundary
 clone_end:5p6
 site:
 end_sequence:BZ170609"

misc_feature
 1. 2063
 /note="wgs_end_extension
 clone_end:77"
 /note="clone_boundary
 clone_end:77"
 site:
 end_sequence:BZ170605"
 complement(195447..196272)
 /note="clone_boundary
 clone_end:5p6
 site:
 end_sequence:BZ170609"

misc_feature
 1. 2063
 /note="wgs_end_extension
 clone_end:77"
 /note="clone_boundary
 clone_end:77"
 site:
 end_sequence:BZ170605"
 complement(195447..196272)
 /note="clone_boundary
 clone_end:5p6
 site:
 end_sequence:BZ170609"

misc_feature
 1. 2063
 /note="wgs_end_extension
 clone_end:77"
 /note="clone_boundary
 clone_end:77"
 site:
 end_sequence:BZ170605"
 complement(195447..196272)
 /note="clone_boundary
 clone_end:5p6
 site:
 end_sequence:BZ170609"

misc_feature
 1. 2063
 /note="wgs_end_extension
 clone_end:77"
 /note="clone_boundary
 clone_end:77"
 site:
 end_sequence:BZ170605"
 complement(195447..196272)
 /note="clone_boundary
 clone_end:5p6
 site:
 end_sequence:BZ170609"

misc_feature
 1. 2063
 /note="wgs_end_extension
 clone_end:77"
 /note="clone_boundary
 clone_end:77"
 site:
 end_sequence:BZ170605"
 complement(195447..196272)
 /note="clone_boundary
 clone_end:5p6
 site:
 end_sequence:BZ170609"

misc_feature
 1. 2063
 /note="wgs_end_extension
 clone_end:77"
 /note="clone_boundary
 clone_end:77"
 site:
 end_sequence:BZ170605"
 complement(195447..196272)
 /note="clone_boundary
 clone_end:5p6
 site:
 end_sequence:BZ170609"

misc_feature
 1. 2063
 /note="wgs_end_extension
 clone_end:77"
 /note="clone_boundary
 clone_end:77"
 site:
 end_sequence:BZ170605"
 complement(195447..196272)
 /note="clone_boundary
 clone_end:5p6
 site:
 end_sequence:BZ170609"

misc_feature
 1. 2063
 /note="wgs_end_extension
 clone_end:77"
 /note="clone_boundary
 clone_end:77"
 site:
 end_sequence:BZ170605"
 complement(195447..196272)
 /note="clone_boundary
 clone_end:5p6
 site:
 end_sequence:BZ170609"

misc_feature
 1. 2063
 /note="wgs_end_extension
 clone_end:77"
 /note="clone_boundary
 clone_end:77"
 site:
 end_sequence:BZ170605"
 complement(195447..196272)
 /note="clone_boundary
 clone_end:5p6
 site:
 end_sequence:BZ170609"

misc_feature
 1. 2063
 /note="wgs_end_extension
 clone_end:77"
 /note="clone_boundary
 clone_end:77"
 site:
 end_sequence:BZ170605"
 complement(195447..196272)
 /note="clone_boundary
 clone_end:5p6
 site:
 end_sequence:BZ170609"

misc_feature
 1. 2063
 /note="wgs_end_extension
 clone_end:77"
 /note="clone_boundary
 clone_end:77"
 site:
 end_sequence:BZ170605"
 complement(195447..196272)
 /note="clone_boundary
 clone_end:5p6
 site:
 end_sequence:BZ170609"

misc_feature
 1. 2063
 /note="wgs_end_extension
 clone_end:77"
 /note="clone_boundary
 clone_end:77"
 site:
 end_sequence:BZ170605"
 complement(195447..196272)
 /note="clone_boundary
 clone_end:5p6
 site:
 end_sequence:BZ170609"

misc_feature
 1. 2063
 /note="wgs_end_extension
 clone_end:77"
 /note="clone_boundary
 clone_end:77"
 site:
 end_sequence:BZ170605"
 complement(195447..196272)
 /note="clone_boundary
 clone_end:5p6
 site:
 end_sequence:BZ170609"

misc_feature
 1. 2063
 /note="wgs_end_extension
 clone_end:77"
 /note="clone_boundary
 clone_end:77"
 site:
 end_sequence:BZ170605"
 complement(195447..196272)
 /note="clone_boundary
 clone_end:5p6
 site:
 end_sequence:BZ170609"

misc_feature
 1. 2063
 /note="wgs_end_extension
 clone_end:77"
 /note="clone_boundary
 clone_end:77"
 site:
 end_sequence:BZ170605"
 complement(195447..196272)
 /note="clone_boundary
 clone_end:5p6
 site:
 end_sequence:BZ170609"

misc_feature
 1. 2063
 /note="wgs_end_extension
 clone_end:77"
 /note="clone_boundary
 clone_end:77"
 site:
 end_sequence:BZ170605"
 complement(195447..196272)
 /note="clone_boundary
 clone_end:5p6
 site:
 end_sequence:BZ170609"

misc_feature
 1. 2063
 /note="wgs_end_extension
 clone_end:77"
 /note="clone_boundary
 clone_end:77"
 site:
 end_sequence:BZ170605"
 complement(195447..196272)
 /note="clone_boundary
 clone_end:5p6
 site:
 end_sequence:BZ170609"

misc_feature
 1. 2063
 /note="wgs_end_extension
 clone_end:77"
 /note="clone_boundary
 clone_end:77"
 site:
 end_sequence:BZ170605"
 complement(195447..196272)
 /note="clone_boundary
 clone_end:5p6
 site:
 end_sequence:BZ170609"

misc_feature
 1. 2063
 /note="wgs_end_extension
 clone_end:77"
 /note="clone_boundary
 clone_end:77"
 site:
 end_sequence:BZ170605"
 complement(195447..196272)
 /note="clone_boundary
 clone_end:5p6
 site:
 end_sequence:BZ170609"

misc_feature
 1. 2063
 /note="wgs_end_extension
 clone_end:77"
 /note="clone_boundary
 clone_end:77"
 site:
 end_sequence:BZ170605"
 complement(195447..196272)
 /note="clone_boundary
 clone_end:5p6
 site:
 end_sequence:BZ170609"

misc_feature
 1. 2063
 /note="wgs_end_extension
 clone_end:77"
 /note="clone_boundary
 clone_end:77"
 site:
 end_sequence:BZ170605"
 complement(195447..196272)
 /note="clone_boundary
 clone_end:5p6
 site:
 end_sequence:BZ170609"

misc_feature
 1. 2063
 /note="wgs_end_extension
 clone_end:77"
 /note="clone_boundary
 clone_end:77"
 site:
 end_sequence:BZ170605"
 complement(195447..196272)
 /note="clone_boundary
 clone_end:5p6
 site:
 end_sequence:BZ170609"

misc_feature
 1. 2063
 /note="wgs_end_extension
 clone_end:77"
 /note="clone_boundary
 clone_end:77"
 site:
 end_sequence:BZ170605"
 complement(195447..196272)
 /note="clone_boundary
 clone_end:5p6
 site:
 end_sequence:BZ170609"

misc_feature
 1. 2063
 /note="wgs_end_extension
 clone_end:77"
 /note="clone_boundary
 clone_end:77"
 site:
 end_sequence:BZ170605"
 complement(195447..196272)
 /note="clone_boundary
 clone_end:5p6
 site:
 end_sequence:BZ170609"

misc_feature
 1. 2063
 /note="wgs_end_extension
 clone_end:77"
 /note="clone_boundary
 clone_end:77"
 site:
 end_sequence:BZ170605"
 complement(195447..196272)
 /note="clone_boundary
 clone_end:5p6
 site:
 end_sequence:BZ170609"

misc_feature
 1. 2063
 /note="wgs_end_extension
 clone_end:77"
 /note="clone_boundary
 clone_end:77"
 site:
 end_sequence:BZ170605"
 complement(195447..196272)
 /note="clone_boundary
 clone_end:5p6
 site:
 end_sequence:BZ170609"

misc_feature
 1. 2063
 /note="wgs_end_extension
 clone_end:77"
 /note="clone_boundary
 clone_end:77"
 site:
 end_sequence:BZ170605"
 complement(195447..196272)
 /note="clone_boundary
 clone_end:5p6
 site:
 end_sequence:BZ170609"

misc_feature
 1. 2063
 /note="wgs_end_extension
 clone_end:77"
 /note="clone_boundary
 clone_end:77"
 site:
 end_sequence:BZ170605"
 complement(195447..196272)
 /note="clone_boundary
 clone_end:5p6
 site:
 end_sequence:BZ170609"

misc_feature
 1. 2063
 /note="wgs_end_extension
 clone_end:77"
 /note="clone_boundary
 clone_end:77"
 site:
 end_sequence:BZ170605"
 complement(195447..196272)
 /note="clone_boundary
 clone_end:5p6
 site:
 end_sequence:BZ170609"

misc_feature
 1. 2063
 /note="wgs_end_extension
 clone_end:77"
 /note="clone_boundary
 clone_end:77"
 site:
 end_sequence:BZ170605"
 complement(195447..196272)
 /note="clone_boundary
 clone_end:5p6
 site:
 end_sequence:BZ170609"

misc_feature
 1. 2063
 /note="wgs_end_extension
 clone_end:77"
 /note="clone_boundary
 clone_end:77"
 site:
 end_sequence:BZ170605"
 complement(195447..196272)
 /note="clone_boundary
 clone_end:5p6
 site:
 end_sequence:BZ170609"

misc_feature
 1. 2063
 /note="wgs_end_extension
 clone_end:77"
 /note="clone_boundary
 clone_end:77"
 site:
 end_sequence:BZ170605"
 complement(195447..196272)
 /note="clone_boundary
 clone_end:5p6
 site:
 end_sequence:BZ170609"

misc_feature
 1. 2063
 /note="wgs_end_extension
 clone_end:77"
 /note="clone_boundary
 clone_end:77"
 site:
 end_sequence:BZ170605"
 complement(195447..196272)
 /note="clone_boundary
 clone_end:5p6
 site:
 end_sequence:BZ170609"

misc_feature
 1. 2063
 /note="wgs_end_extension
 clone_end:77"
 /note="clone_boundary
 clone_end:77"
 site:
 end_sequence:BZ170605"
 complement(195447..196272)
 /note="clone_boundary
 clone_end:5p6
 site:
 end_sequence:BZ170609"

misc_feature
 1. 2063
 /note="wgs_end_extension
 clone_end:77"
 /

Project Information
Center Project Name: 1902639
Center clone name: RPCI-23_431K23

Summary Statistics

Consensus quality: 193161 bases at least Q40
Consensus quality: 201982 bases at least Q30
Consensus quality: 204342 bases at least Q20
Estimated insert size: 190000; agarose-gel estimation
Estimated insert size: 212596; sum-of-contigs estimation
Quality coverage: 8.52 in Q20 bases; agarose-gel estimation
Quality coverage: 7.62 in Q20 bases; sum-of-contigs estimation.

NOTE: This is a 'working draft' sequence. It currently consists of 27 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1 1535: contig of 1535 bp in length
1536 1635: gap of unknown length
1636 2881: contig of 1246 bp in length
2882 2981: gap of unknown length
2982 4237: contig of 1256 bp in length
4238 4337: gap of unknown length
4338 5459: contig of 1122 bp in length
5460 5559: gap of unknown length
5560 6642: contig of 1083 bp in length
6643 6743: gap of unknown length
6743 8140: contig of 1398 bp in length
8141 8240: gap of unknown length
8241 10071: contig of 1831 bp in length
10072 10171: gap of unknown length
10172 11363: contig of 1192 bp in length
11364 11463: gap of unknown length
11464 14217: contig of 2754 bp in length
14218 14317: gap of unknown length
14318 17815: contig of 3498 bp in length
17816 17915: gap of unknown length
17916 20634: contig of 2719 bp in length
20635 20734: gap of unknown length
20735 24318: contig of 3584 bp in length
24319 24418: gap of unknown length
24419 27634: contig of 3216 bp in length
27635 27734: gap of unknown length
27735 32614: contig of 4880 bp in length
32615 32714: gap of unknown length
32715 37575: contig of 4865 bp in length
37576 37675: gap of unknown length
37676 44946: contig of 7267 bp in length
44947 45046: gap of unknown length
45047 50203: contig of 5157 bp in length
50204 50303: gap of unknown length
50304 54763: contig of 4460 bp in length
54764 54863: gap of unknown length
54864 65724: contig of 10861 bp in length
65725 65824: gap of unknown length
65825 75736: contig of 9912 bp in length
75737 75836: gap of unknown length
75837 84975: contig of 9139 bp in length
84976 85075: gap of unknown length
85076 101226: contig of 16151 bp in length
101227 101326: gap of unknown length
101327 120390: contig of 19064 bp in length
120391 120490: gap of unknown length
120491 135635: contig of 15145 bp in length
135636 135735: gap of unknown length
135736 162383: contig of 26648 bp in length
162384 162483: gap of unknown length
162484 188312: contig of 25829 bp in length
188313 188412: gap of unknown length

FEATURES

Source

* 188413 215196: contig of 26784 bp in length.
Location/Qualifiers
1..215196
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/clone="RP23-431K23"
/clone_11b="RPCI mouse BAC library 23"

ORIGIN

Query Match 87.0%; Score 17.4; DB 2; Length 215196;
Best Local Similarity 94.7%; Pred. No. 1.4e+02;
Matches 18; Conserved 0; Mismatches 1; Indels 0; Gaps 0;

CY 1 AGTCCAAACGATGACATG 19
|||
Db 213819 AGTCCAAACGATGACATG 213837

RESULT 10

AC149052

LOCUS AC149052 215199 bp DNA linear HTG 04-AUG-2004
DEFINITION Mus musculus chromosome 7 clone RP24-54801, WORKING DRAFT SEQUENCE,
8 unordered pieces.

ACCESSION AC149052
KEYWORDS HTG; HTGS PHASE1; HTGS DRAFT; HTGS_ACTIVEFIN.

SOURCE

Mus musculus (house mouse)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathia; Muridae; Murinae; Mus.

REFERENCE

1 (bases 1 to 215199)

AUTHORS

Wilson, R.K.

TITLE

The sequence of Mus musculus clone

JOURNAL

Unpublished

AUTHORS

Wilson, R.K.

TITLE

Submitted (08-MAY-2004) Genome Sequencing Center, 4444 Forest Park

JOURNAL

3 (bases 1 to 215199)

REFERENCE

Submitted (04-AUG-2004) Genome Sequencing Center, 4444 Forest Park

AUTHORS

Wilson, R.K.

TITLE

Direct Submission

JOURNAL

On Aug 4, 2004 this sequence version replaced gi:47084635.

COMMENT

----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu

Contact: submissions@wustl.edu
----- Project Information -----
Center project name: M_B80548001

----- Summary Statistics -----
Sequencing vector: M13; 0%

Chemistry: Dye-terminator Big Dye; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990319

Consensus quality: 210989 bases at least Q40
Consensus quality: 211541 bases at least Q30
Consensus quality: 211780 bases at least Q20

NOTE: This is a 'working draft' sequence. It currently consists of 8 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.


```

* 1 1024: contig of 1024 bp in length
* 1025 1124: gap of unknown length
* 1125 2558: contig of 1434 bp in length
* 2559 2658: gap of unknown length
* 2659 3852: contig of 1194 bp in length
* 3853 23853: gap of unknown length
* 23854 23953: contig of 19901 bp in length
* 23954 46296: gap of unknown length
* 46297 46396: contig of 22343 bp in length
* 46397 96500: gap of unknown length
* 96501 96601: contig of 50104 bp in length
* 96601 152172: gap of unknown length
* 152173 152272: contig of 55572 bp in length
* 152273 215199: gap of unknown length
* 215199 62927: contig of 62927 bp in length.
Location/Qualifiers
1. 215199
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="7"
/clone="RP24-54801"
1. 1024
/note="assembly_name:Contig12"
1125. 2558
/note="assembly_name:Contig16"
2559. 3852
/note="assembly_name:Contig27"
3953. 123853
/note="assembly_name:Contig36"
23954. 46296
/note="assembly_name:Contig37"
clone_end:T7
vector_side:right"
46397. 96500
/note="assembly_name:Contig38"
96601. 152172
/note="assembly_name:Contig39"
152273. 215199
/note="assembly_name:Contig40"

ORIGIN
Query Match 87.0%: Score 17.4; DB 2; Length 215199;
Best Local Similarity 94.7%: Pired. No. 1.4e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 AGTCGCAACGATACATG 19
Db 113302 AGTCGCAACGATACATG 113320

RESULT 11
AC140735 229472 bp DNA linear HTG 27-MAR-2003
LOCUS Rattus norvegicus clone CH230-199G5, WORKING DRAFT SEQUENCE, 50
DEFINITION Unordered pieces.
ACCESSION AC140735 GI:28894512
VERSION AC140735.4 GI:28894512
KEYWORDS HTG; HTGS PHASE1; HTGS DRAFT.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 229472)
Muzny,D,Marler,M,Mezker,M,Lee,S,Abramson,S,Adams,C,Alder,J,
Allen,C,Allen,H,Aisbrook,S,Amin,A,Anguiano,D,
Aryalambuchi,V,Aoyagi,A,Ayodeji,M,Baca,E,Baden,H,
Baldwin,D,Bandaranaike,D,Barber,M,Barnstead,M,Benahmed,F,
Biswal,R,Bisai,J,Blankenburg,K,Blyth,P,Brown,M,
Bryant,N,Buhay,C,Burch,P,Burrell,K,Calderson,E,
Cardenas,V,Carter,K,Cavazos,I,Ceasar,H,Center,A,
Chacko,J,Chavez,D,Chen,G,Chen,R,Chen,Y,Chen,Z,Chu,J,

```

```

Cleveland,C,Cockrell,R,Cox,C,Coyle,M,Cree,A,D'Souza,L,
Davila,M,L,Davis,C,Davy-Carroll,L,De Anda,C,Dederich,D,
Delgado,O,Denson,S,Derramo,C,Ding,Y,Dinh,H,Divya,K,
Diaper,H,Dugan-Rocha,S,Dunn,A,Durbin,K,Duval,B,Eaves,K,
Egan,A,Escotto,M,Eugene,C,Evans,C,A,Falls,T,Fan,G,
Fernandez,S,Finley,M,Flagg,N,Folbes,L,Foster,M,Foster,P,
Fraser,C,M,Gabisi,A,Ganta,R,Garcia,A,Garner,T,Garza,M,
Georgievski,B,Geer,R,Gill,R,Grady,M,Guerra,M,Guevara,W,
Gunaratne,P,Haaland,M,Hamil,C,Hamilton,C,Hamilton,K,
Harvey,Y,Havlik,P,Hawes,A,Henderson,N,Hernandez,J,
Hernandez,R,Hines,S,Hladun,S,L Hodgson,A,Hogues,M,
Hollins,B,Howell,S,Hulky,S,Hume,J,Idlebird,D,Jackson,A,
Jackson,S,Jacob,L,Jiang,H,Johnson,B,Johnson,R,Jolivet,A,
Karpachy,S,Kelly,S,Khan,Z,King,L,Kovar,C,
Kowals,C,Kraft,C,L,Lebow,H,Levan,D,Lewis,L,Li,Z,Liu,J,
Liu,J,Liu,W,Liu,Y,London,P,Longacre,S,Lopez,J,
Lorenschweta,L,Louiseged,H,Lozado,R,J,Lu,X,Ma,J,
Maheshwari,M,Mahindaratne,M,Mahmoud,M,Mallory,K,Mangum,A,
Mangum,B,Mapua,P,Martin,K,Martin,R,Martinez,E,
Mawhinney,S,McLeod,M,McNeill,T,Meenen,E,Milosavljevic,A,
Minier,G,Mirja,E,Montemayor,U,Morre,S,Morgan,M,Morris,K,
Morris,S,Kumidasa,M,Murphy,M,Nair,L,Nankervis,C,Neal,D,
Newton,N,Nguyen,N,Norris,S,Nwokelemeh,O,Okwonon,G,
Olatunmbunsegun,A,Pal,S,Parks,K,Pasternak,S,Paul,H,
Perez,A,Perez,L,Pfannkuch,C,Plopper,F,Poindexter,A,
Popovic,D,Primus,B,Pu,L-L,Puazo,M,Quiroz,J,Rachlin,E,
Reeves,K,Regier,M,A,Reigh,R,Reilly,B,Reilly,M,Reyn,J,
Reuter,M,Richards,S,Riggs,F,Rives,C,Rodkey,T,Rojas,A,
Rose,M,Rose,R,Ruiz,S,J,Sanders,W,Savery,G,Scherer,S,
Scott,G,Shatman,S,Shen,H,Shelly,J,Shvartsbeyn,A,
Sisson,I,Sitter,C,D,Smajd,D,Speed,A,Sodergren,E,
Song,X-Z,Sorrell,R,Sosa,J,Steime,M,Strong,R,Sutton,A,
Svatek,A,Tabor,P,Taylor,C,Taylor,T,Thomas,N,Thomas,S,
Tingey,A,Trejos,Z,Usmami,K,Valas,R,Vera,V,Villaana,D,
Waldron,L,Walker,B,Wang,U,Wang,Q,Wang,S,Warren,J,
Warren,R,Wel,X,White,F,Williams,G,Willson,K,Wieczyk,R,
Wooten,H,Worley,K,Wright,D,Wright,R,Wu,J,Yakub,X,
Yen,J,Yoon,L,Yoon,V,Yu,F,Zhang,Y,Zhou,Z,Zhou,X,
Zhao,S,Zunn,D,von Niederhausern,A,Weiss,R,Smith,D,R,
Holt,R,A,Smith,H,O,Weinstock,G, and Gibbs,R,A.

Direct Submission
Unpublished
2 (bases 1 to 229472)
Worley,K,C.
Direct Submission
Submitted (01-MAR-2003) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 229472)
Worley,K,C.
Direct Submission
Submitted (27-MAR-2003) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Mar 10, 2003 this sequence version replaced gi:28893586.

----- genome center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: KEBB
Center clone name: CH230-199G5
----- Summary Statistics
Sequencing vector: Plasmid;
Chemistry: Dye-terminator Big Dye 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 213297 bases at least Q40
Consensus quality: 218310 bases at least Q30
Consensus quality: 22517 bases at least Q20
Estimated insert size: 211997; sum-of-contigs estimation
Quality coverage: 4x in Q20 bases; sum-of-contigs estimation

```

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 50 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1
1171: contig of 1171 bp in length
1172: gap of unknown length
1271: contig of 1065 bp in length
2336: gap of unknown length
2436: contig of 1067 bp in length
3503: gap of unknown length
3603: contig of 1024 bp in length
4627: gap of unknown length
4727: contig of 1234 bp in length
5961: gap of unknown length
6061: contig of 1106 bp in length
7167: gap of unknown length
7267: contig of 1496 bp in length
8763: gap of unknown length
8764: contig of 1141 bp in length
10004: gap of unknown length
10104: contig of 1515 bp in length
11629: gap of unknown length
11719: contig of 1789 bp in length
13508: gap of unknown length
13608: contig of 1580 bp in length
15188: gap of unknown length
15288: contig of 1677 bp in length
15965: gap of unknown length
17065: contig of 1203 bp in length
18268: gap of unknown length
18368: contig of 1959 bp in length
20327: gap of unknown length
20427: contig of 1886 bp in length
22313: gap of unknown length
22413: contig of 1167 bp in length
23580: gap of unknown length
23680: contig of 1346 bp in length
25026: gap of unknown length
25126: contig of 1683 bp in length
25127: gap of unknown length
26809: contig of 1703 bp in length
26909: gap of unknown length
28612: contig of 1703 bp in length
28712: gap of unknown length
28713: contig of 1683 bp in length
30395: gap of unknown length
30495: contig of 1377 bp in length
31872: gap of unknown length
31972: contig of 1287 bp in length
31973: gap of unknown length
33259: contig of 2073 bp in length
33359: gap of unknown length
35432: contig of 1436 bp in length
35532: gap of unknown length
36968: contig of 1276 bp in length
37068: gap of unknown length
38344: contig of 1567 bp in length
38445: gap of unknown length
40011: contig of 1036 bp in length
40112: gap of unknown length
41147: contig of 2091 bp in length
41247: gap of unknown length
43338: contig of 2005 bp in length
43438: gap of unknown length
45443: contig of 2565 bp in length
45543: gap of unknown length
48112: contig of 2383 bp in length
48212: gap of unknown length
50595: contig of 1770 bp in length
50695: gap of unknown length
52465: contig of 1770 bp in length

52466: gap of unknown length
52566: contig of 2978 bp in length
55543: gap of unknown length
55544: contig of 2878 bp in length
58521: gap of unknown length
58522: contig of 2429 bp in length
61051: gap of unknown length
61151: contig of 3108 bp in length
64258: gap of unknown length
64359: contig of 4171 bp in length
68529: gap of unknown length
68629: contig of 4079 bp in length
72709: gap of unknown length
72809: contig of 7031 bp in length
79839: gap of unknown length
79940: contig of 7074 bp in length
79940: gap of unknown length
87014: contig of 6009 bp in length
87114: gap of unknown length
93123: contig of 7290 bp in length
93223: gap of unknown length
100513: contig of 8553 bp in length
100613: gap of unknown length
109171: contig of 9618 bp in length
109172: gap of unknown length
109272: contig of 9618 bp in length
118890: gap of unknown length
118989: contig of 13063 bp in length
132053: gap of unknown length
132153: contig of 10557 bp in length
142709: gap of unknown length
142809: contig of 17266 bp in length
142810: gap of unknown length
160095: contig of 16878 bp in length
160196: gap of unknown length
177074: contig of 22498 bp in length
177173: gap of unknown length
199671: contig of 22498 bp in length

Query Match
Best Local Similarity 94.7%; Pred. No. 1,4e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GTGGCAACAGTACATGCG 20
Db 18905 GTGGCAACAGTACATGCG 18923

RESULT 12
AC136831/c
LOCUS
DEFINITION
Rattus norvegicus clone CH230-7418, WORKING DRAFT SEQUENCE, 3
AC136831 243210 bp DNA linear HTG 23-NOV-2002
unorderd pieces.
AC136831
VERSION
AC136831.2 GI:25188459
HTG; HTGS PHASE1; HTGS DRAFT; HTGS_FULFILLTOP.
SOURCE
Rattus norvegicus (Norway rat)
ORGANISM
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Scuriognathia; Muridae; Murinae;
Rattus.
1 (bases 1 to 243210)
Munry,D,Marie, Mezker,M,Lee, Abramson,S, Adams,C, Alder,J,
Allen,C, Allen,H, Alsbrooks,S, Amin,A, Anguiano,D,
Anyalebechi,V, Aoyagi,A, Ayodeji,M, Baca,E, Baden,H,
Baldwin,D, Bandaranaike,D, Barber,M, Barnstead,M, Benahmed,F,
Biswalto,K, Blair,J, Blankenburg,K, Blyth,P, Brown,M,
Bryant,N, Buhay,C, Burch,P, Surrell,K, Galderson,E,
Gardenas,V, Garter,K, Cavazos,I, Ceasar,H, Center,A,
Chacko,J, Chavez,D, Chen,G, Chen,R, Chen,Y, Chen,Z, Chu,J,
Cleveland,C, Cockrell,R, Cox,C, Coyte,M, Cree,A, D'Souza,L,
Davila,M,L, Davis,C, Davy-Carroll,L, De Anda,C, Dedereich,D,
Delgado,O, Denson,S, Deramo,C, Ding,Y, Dinh,H, DiVya,K,
Draper,H, Dugan-Rocha,S, Dunn,A, Durbin,K, Duval,B, Eaves,K,
Egan,A, Escotto,X, Eugene,C, Evans,C,A, Falls,T, Fan,G,
Fernandez,S, Finley,M, Flagg,N, Forbes,L, Foster,M, Foster,P,
Fraser,C,M, Gabisi,A, Ganta,R, Garcia,A, Garner,T, Garza,M,

Gebreygeorgis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W., Gunaratne, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K., Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S.L., Hodgson, A., Hogues, M., Hollins, B., Howells, S., Hulky, S., Hume, J., Idlebird, D., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Joliver, A., Karpach, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kowis, C., Kraft, C.L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lorenz, L., Louised, L., Lozada, R.J., Lu, X., Ma, J., Maheshwari, M., Mahindaratne, M., Mahmoud, M., Mallory, K., Mangum, A., Mangum, B., Mapa, P., Martin, K., Martin, R., Martinez, E., Mawhinney, S., McLeod, M.P., McNeill, T.Z., Meenen, E., Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Mundasa, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nsoekeme, O., Okunolu, G., Olarnpunsagoon, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfankuch, C., Plopper, F., Polidexter, A., Popovic, D., Primus, E., Pt, L., Plazo, M., Quiroz, J., Rachtin, E., Reeves, K., Regier, M.A., Reigh, R., Kelly, B., Kelly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rives, C., Rodkey, T., Rojase, A., Rose, M., Rose, R., Ruiz, S.J., Sanders, W., Savary, G., Scherer, S., Scott, G., Shatsman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C.D., Smajls, D., Sneed, A., Sodergren, E., Song, X.-Z., Sorrelle, R., Soza, J., Steinle, M., Strong, R., Sutton, A., Syatek, A., Taber, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K., Valas, R., Vera, V., Villaseña, D., Waldron, J., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wleczky, R., Woodson, H., Wortley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, Y., Yoon, L., Yoon, V., Yu, P., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Niederhausen, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O., Weinstock, G., and Gibbs, R.A.

Direct Submission
Unpublished
2 (bases 1 to 243210)
Rat Genome Sequencing Consortium.
Submitted (08-NOV-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 243210)
Rat Genome Sequencing Consortium.
Direct Submission
Submitted (23-NOV-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
On Nov 23, 2002 this sequence version replaced gi:24796705.
The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a contig-scaffold). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GTOT
Center clone name: CH230-7418
----- Summary Statistics
Assembly program: Phrap; version 0.990329
Consensus quality: 230829 bases at least Q40
Consensus quality: 233240 bases at least Q30

Consensus quality: 234747 bases at least Q20
Estimated insert size: 236645; sum-of-contigs estimation
Quality coverage: 7x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a working draft sequence. It currently
* consists of 3 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 240139: contig of 240139 bp in length
* 240140 240239: gap of unknown length
* 240240 241508: contig of 1269 bp in length
* 241509 241608: gap of unknown length
* 241609 243210: contig of 1602 bp in length.
* Location/Qualifiers
* 1. 243210
* /organism="Rattus norvegicus"
* /mol_type="genomic DNA"
* /db_xref="taxon:10116"
* /clone="CH230-7418"
* 1. 1466
* /note="wgs_contig"

misc_feature
ORIGIN
Query Match 87.0%; Score 17.4; DB 2; Length 243210;
Best Local Similarity 94.7%; Pred. No. 1.4e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0;
Gaps 0;

QY 1 AGTCGCAACAGTACATG 19
Db 74206 AGACCGCAACAGTACATG 74188

RESULT 13
AC115331
LOCUS
DEFINITION
Rattus norvegicus clone CH230-307019. *** SEQUENCING IN PROGRESS
** 7 unordered pieces.
AC115331
AC115331.3 GI:23270141
HTG: HTGS_PHASE1: HTGS_DRAFT; HTGS_ENRICHED.
Rattus norvegicus (Norway rat)
SOURCE
ORGANISM
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 243702)
1 Muzny, D.Marie, Metzker, M.Lee, Abramson, S., Adams, C., Alder, J.,
Allen, C., Allen, H., Altschrocks, S., Amin, A., Angiano, D.,
Anyalebechi, V., Aoyagi, A., Ayodeji, W., Bacc, E., Baden, H.,
Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Barnham, F.,
Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M.,
Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E.,
Cardenas, V., Carter, K., Cavazos, I., Caesar, H., Center, A.,
Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J.,
Cleveland, C., Cockrell, R., Cox, C., Coyle, K., Cree, A., D'Souza, L.,
Davila, M.L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D.,
Delgado, O., Denison, S., Deramo, C., Ding, Y., Dinh, H., Divya, K.,
Diaper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duvall, B., Eaves, K.,
Egan, A., Escotto, M., Eugene, C., Evans, C.A., Falls, T., Fan, G.,
Fernandez, S., Flinkey, N., Flagg, N., Forbes, J., Foster, M., Foster, P.,
Fraser, C.M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M.,
Gebreygeorgis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W.,
Gunaratne, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K.,
Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J.,
Hernandez, R., Hines, S., Hladun, S.L., Hodgson, A., Hogues, M.,
Hollins, B., Howells, S., Hulky, S., Hume, J., Idlebird, D., Jackson, A.,
Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Joliver, A.,

Karpachy, S., Kelly, S., Khan, Z., King, L., Kovari, C., Kowitz, C., Kraft, C.L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lorenshewa, L., Louisedge, H., Lozano, R.J., Lu, X., Ma, J., Maheshwari, M., Mahindaratne, M., Mahmoud, M., Mallory, K., Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E., Mawhney, S., McLeod, M.P., McNeill, T.Z., Meenen, E., Mijosevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nait, L., Nankervis, C., Neal, D., Newton, N., Ngunda, N., Norris, S., Nsoekeme, O., Okunolu, G., Olarnunsoo, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfankuch, C., Plopper, F., Poindexter, A., Popovic, D., Primm, E., Pu, L., Puzo, M., Quirroz, J., Rachlin, E., Reeves, K., Regier, M.A., Reich, R., Reilly, B., Reilly, M., Ren, Y., Reuter, R., Richards, S., Riggs, F., Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S., Sanders, W., Saverly, G., Scherer, S., Scott, G., Shatsman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C.D., Smaj, D., Sneed, A., Sodergren, E., Song, X.-Z., Sorrelle, R., Soga, J., Steinle, M., Strong, R., Sutton, A., Syarik, A., Taber, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Uman, K., Valas, R., Vera, V., Villalana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wleczky, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, Y., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Zhao, S., Dunn, D., von Niederhausen, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O., Weinstock, G., and Gibbs, R.A.

 TITLE Direct Submission
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 243702)
 AUTHORS Morley, K.C.
 TITLE Direct Submission
 JOURNAL Submitted (17-MAR-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
 3 (bases 1 to 243702)
 Rat Genome Sequencing Consortium.
 Direct Submission
 Submitted (23-SEP-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
 On Sep 23, 2002 this sequence version replaced gi:21781813.
 The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/atlas/>). As a result, the sequence may extend beyond the ends of the contigs and there may be contigs that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center

 Center: Baylor College of Medicine
 Center code: BCM
 Web site: <http://www.hgsc.bcm.tmc.edu/>
 Contact: hgsc-help@bcm.tmc.edu
 Project Information
 Project name: G00A
 Center project name: G00A
 Center clone name: CH230-307D19
 ----- Summary Statistics
 Assembly program: Phrap; version 0.990329
 Consensus quality: 175747 bases at least Q40
 Consensus quality: 178942 bases at least Q30
 Consensus quality: 180494 bases at least Q20
 Estimated insert size: 199396; sum-of-contigs estimation
 Quality coverage: 4x in Q20 bases; sum-of-contigs estimation

 NOTE: Estimated insert size may differ from sequence length
 * (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 7 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.

***** This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

1 3679: contig of 3679 bp in length
 * 3680 3779: gap of unknown length
 * 3780 208206: contig of 204427 bp in length
 * 208307 208306: gap of unknown length
 * 208307 224315: contig of 26009 bp in length
 * 224316 224315: gap of unknown length
 * 224316 235574: contig of 1159 bp in length
 * 235575 235674: gap of unknown length
 * 235675 237230: contig of 1556 bp in length
 * 237231 237330: gap of unknown length
 * 237331 237330: gap of unknown length
 * 237331 239464: contig of 2134 bp in length
 * 239465 239564: gap of unknown length
 * 239565 243702: contig of 4138 bp in length.

Location/Qualifiers
 1. 243702
 /organism="Rattus norvegicus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10116"
 /clone="CH230-307D19"
 1. 1338
 /note="wgs_contig"
 3780. 4825
 /note="wgs_contig"
 206630. 208206
 /note="wgs_contig"

ORIGIN
 Query Match 87.0%; Score 17.4; DB 2; Length 243702;
 Best Local Similarity 94.7%; Pred. No. 1,4e+02;
 Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

CY 2 GTCCCAACAGTACATCG 20
 Db 223467 GTGGCAACAGTACATCG 223485

RESULT 14
 AC094556/c
 LOCUS AC094556
 DEFINITION Rattus norvegicus clone CH230-3P4, WORKING DRAFT SEQUENCE, 5
 UNORDERED PAGES.
 AC094556
 AC094556.6 GI:30466689
 HTG; HTGS PHASE1; HTGS DRAFT; HTGS FULLTOP.
 KEYWORDS
 Rattus norvegicus (Norway rat)
 SOURCE
 ORGANISM
 Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.
 1 (bases 1 to 289856)
 Muray, D., Marle, M., Mezzker, M., Lee, A., Abramson, S., Adams, C., Alder, J., Allen, C., Allen, H., Albrooke, S., Amin, A., Anguiano, D., Anyalebechi, V., Ayoyagi, A., Ayodeji, M., Baca, E., Baden, H., Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F., Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E., Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Chen, A., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Cleveland, C., Cockrell, R., Cox, C., Coyly, M., Cree, A., D'Souza, L., Davila, M., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D., Delgado, O., Denson, S., Detaro, C., Ding, Y., Dinh, H., Divya, K., Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K., Egan, A., Escotto, M., Eugene, C., Evans, C.A., Falls, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, J., Foster, M., Foster, P., Fraser, C.M., Gabisi, A., Gante, R., Garcia, A., Garner, T., Garza, M., Gebregeorgis, E., Geer, K., Gill, R., Grady, M., Guerra, M., Guevara, W., Gunaratne, P., Haaland, W., Hamil, C., Hamilton, C., Hamilton, K., Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S.L., Hodgson, A., Hogues, M., Hollins, B., Howells, S., Huliy, S., Hume, J., Idelbird, D., Jackson, A.,

Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A., Karpeth, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kowit, C., Kraft, C.L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, D., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lorenz, L., Lougheed, H., Lozano, R.J., Lu, X., Ma, J., Maheshwari, M., Mahindratne, M., Mahmoud, M., Malloy, K., Mangum, A., Mangum, B., Mapa, P., Martin, K., Martin, R., Martinez, E., Mawhney, S., McLeod, M.P., McNeill, T.Z., Meenen, E., Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Mundasa, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, G., Olanupusoon, A., Pal, S., Paks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Piamkotch, C., Plopper, F., Polindexter, A., Popovic, D., Primus, E., Pu, L.-L., Puzo, M., Quito, J., Rachlin, E., Reeves, K., Regier, M.A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S.J., Sanders, W., Savary, G., Scherer, S., Scott, G., Shatsman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sisson, T., Sitter, C.D., Smajls, D., Sneed, A., Sodergren, E., Song, X.-Z., Sorrelle, R., Soza, J., Steinle, M., Strong, R., Sutton, A., Svatek, A., Taber, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usman, K., Valas, R., Vera, V., Villaseana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, P., Williams, G., Willson, R., Wleczky, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Niederhausen, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O., Weinstock, G., and Gibbs, R.A.

TITLE JOURNAL AUTHORS REFERENCE JOURNAL

Unpublished
2 (bases 1 to 289856)
Direct Submission
Submitted (15-SEP-2001) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 289856)
Rat Genome Sequencing Consortium.

REFERENCE TITLE AUTHORS JOURNAL

COMMENT

Submitted (09-MAY-2003) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
On May 9, 2003 this sequence version replaced gi:24942102.
The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GAWR
Center clone name: CH230-3P4
----- Summary Statistics

Assembly program: Atlas:
Consensus quality: 238163 bases at least Q40
Consensus quality: 238828 bases at least Q30
Consensus quality: 241498 bases at least Q20
Estimated insert size: 246461; sum-of-contigs estimation
Quality coverage: 7x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length

(see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently consists of 5 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. * This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.
* 1 282464: contig of 282464 bp in length
* 282465 282564: gap of unknown length
* 282565 283807: contig of 1243 bp in length
* 283808 283907: gap of unknown length
* 283908 286835: contig of 2928 bp in length
* 286836 286935: gap of unknown length
* 286936 288319: contig of 1384 bp in length
* 288320 288419: gap of unknown length
* 288420 289856: contig of 1437 bp in length.

FEATURES
source
1..289856
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-3P4"
misc_feature
1..2521
/note="wgs contig"
misc_feature
108311..110551
/note="wgs_contig"

Query Match 87.0%; Score 17.4; DB 2; Length 289856;
Best Local Similarity 94.7%; Pred. No. 1.4e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 GTCCGAACAGTACATCG 20
Db 95008 GTGGCAACAGTACATCG 94990

RESULT 15
HHV6U1102 10601 bp DNA linear VRL 27-OCT-1995
LOCUS Human herpesvirus type 6A tp/cap, mdbp, U42orf, U43orf, U44orf, and dUTPase genes.
DEFINITION
X92436
ACCESSION X92436.1 GI:1044869
VERSION dUTPase; mdbp gene; tp/cap gene; U40; U41; U42; U43; U44; U45.
KEYWORDS Human herpesvirus 6
SOURCE Human herpesvirus 6
ORGANISM
viruses; dsDNA viruses, no RNA stage; Herpesviridae; Betaherpesvirinae; Roseolovirus.

REFERENCE
1 Jones, M.D.
DNA sequence analysis of the origin of DNA replication region of Human Herpesvirus 6A strain U1102
Unpublished
2 (bases 1 to 10601)
JONES, M.D.
Direct Submission
Submitted (18-OCT-1995) M.D. Jones, Royal Postgraduate Medical School, Virology, Du Cane Road, London, W12 0NN, UK
Location/Qualifiers
1..10601
/organism="Human herpesvirus 6"
/mol_type="genomic DNA"
/strain="U1102"
/isolate="U1102"
/db_xref="taxon:10368"
/cell_line="T cell"
/lab_host="human T-cells"
complement(1..278)
/gene="tp/cap"
complement(1..278)

FEATURES
source
1..10601
/organism="Human herpesvirus 6"
/mol_type="genomic DNA"
/strain="U1102"
/isolate="U1102"
/db_xref="taxon:10368"
/cell_line="T cell"
/lab_host="human T-cells"
complement(1..278)
/gene="tp/cap"
complement(1..278)

```
/gene="tp/cap"
/feature="U40orf"
/codon_start=1
/protein_id="CAA63166.1"
/db_xref="GI:1044870"
/db_xref="GOA:P52384"
/db_xref="Swiss-Prot:P52384"
/translation="MNSLSQSLVCAHNECALDLEQLKFCDEPVILSMANFKKNGI
VILHLYOTFEFGIKKQMLCSALTYMVLKAMYEQVILLDALES"
/complement(286..3684)
/gene="mbdp"
/complement(286..3684)
/feature="U41orf"
/codon_start=1
/protein_id="CAA63167.1"
/db_xref="GI:1044871"
/db_xref="GOA:P52338"
/db_xref="Swiss-Prot:P52338"
/translation="MADNETTVASAPVSTAAVTPPKKELLDTLSVLSMERNSPI
VISPLMNTLVENDSTTVKTPITNFGGTLITKITSFVYVCFPHGTEBLYGMADHG
DLRLCQTRKPHLOSFEVPTARKVIDIPLALCSAGKADSVTCLGKSGPKELF
AGLIPCEBQIQVQGEVSCVKIPLYSATLFTETETISLSCTEFIQERGFPLAS
ETLFFVFTSMGTTLFRSNTKELIDAGLKQFTDQGTQVLAHPKTYLGIQOKLSAV
EKDFIMLVDSVTLSFSHVAEYLDVDPQIMNENMDPIRNSETHAERMAOLNL
KLHSLHVALIPAPNSILYCSKLAFPNVKQAFNVMTOELLRELSTFCAALSLSD
DYANDKXIKCDSTSGDKDKFSANHLAYACATSPQLSPVVMNMRMAGVANAHT
EYNHLYNCANLCERCDKCCSCIGTAMVRYGRLPAIPKVKKEPILVMESRY
AEVDILSGFRKFSVSEKELIGKQNTLSLDRKTVSQIFDYCKNSLIDPTGEBT
NVRSKDFVSIIGLTQCLIECVSRCTIVEMRTQREQIENCLOSFPNDTTPVAAF
SPPLFSYVAVILTLQNLALIVASGVVDRTCTGNSIKRWLQOYQSYGFHSYL
KKGILNRTAVKVASNMEOIILDCDLYKSGKYKTTIQAOLCLSMOCLDRIKRKP
FNKSKAHNPVFNKVKHKNPLSGCISFLFKYHDLFPVKISLCTELORFLN
NVKRLDIGNDEEVKTRIKAPSLITNTYDEIDILIDPCLISFIDCYKHNKFLSAG
PHDYLHGLTSLKLTQNPVLPVYLIDKQPKSSIQEVLVYKXUVDGYNPIYAS
LSKBPNGTITFSRSLVTGTLLEKFSLANRKYFOGQJMHGSGGVVDNPPVISA
LODFERROKTIITATKFEVITKVRREALIMTEVEXKVALSIVENLINDIPELLI
IAEVMRREDKPTMDMLFVVDGRALASIMKLNHLVDMVVKDFISINLOSVEVY
SSNDAPYDSSEILAEEDQNGVLCDETEETEDPMTKARL"
/complement(3910..3932)
/feature="U42orf"
/codon_start=1
/protein_id="CAA63168.1"
/db_xref="GI:1044872"
/db_xref="GOA:P52354"
/db_xref="Swiss-Prot:P52354"
/translation="MYPRGYKSHDHYKOTAFRTIKSSTHROTSTKSHAKNFRK
LALPKLIDSRDLALSITELKQKTIIEKQOKRANALITFLNLPVPPADPNRS
LKSGLKPTINGDADRPQRILDRFPQTHERICLINDILJTLDLKRYCDSLTRRS
NQOLMDGRIPLSTNSPCLAFVATVEASGYEFHDLNLPVNPDLFENVITTVKPE
FNNKLNAXLTCVNDNGHDIERYKRLQICGXPVLDPMNSELEVOQDTSFRPH
IOQAMSIVTFARILROIKEOIQTKKPFIRIDGRVAEYEGCGLMILGXOFNS
HKDDVSCONRIRIAPWPKSLFCTYAKAPKRLFPNPEEYRNLSTFCPADVT
EPSCYSTNHDLPOTSHRSHKNGTPKVKSKYCKVEKDPDLSITTKTTEILLSEME
TDKTIKPRRLNNAQKEIRIVININENVNSGHSBESSVEMDLIDYALDTCETNIN
ACSSDS"
/complement(6690..9472)
/gene="U43orf"
```

```
CDS
complement(6690..9472)
/gene="U43orf"
/codon_start=1
/protein_id="CAA63169.1"
/db_xref="GI:1044873"
/db_xref="GOA:P52467"
/db_xref="Swiss-Prot:P52467"
/translation="MTIVFATREYDANVLPSTLCRSPSEHLIPITVYKPPNNNSF
LQOTQCKNSRSDTVFVCAEKUNLSHYIQTASPIAEVANSNDKTESLIVDMY
LSQTKERDEKPEKMAVFKSLIKYTKFLPSPFPWFLSTYQGTESMLITMVFY
YLFEOKSITITTKVQVQCFTENGMVPYSSMSERINITLKSRRKLPADSTYAR
QKNLRDEEFGHLDTOINLPKSHLNTFRVHYIAYNTALETTKFFVYCNLTSD
SNLPIGOCCORNLHILGNSLHNLCLMKQYFADVCYKTYIDIKLKNLNAYTEB
YALYSKXKTIQALPSBOTRAIAKCNNGEGLSPVKGIOGLLEISASRYOIODK
RIRYRQHLKDYHRRPVPYVQVLLHKRITPCGNSBDEYEMNGFNRIQIYPRXISDE
ALTRAIALQDTHFLCDDEKOFITTRHEITNERIPVTWIGIDLPLQDTATITBEF
FSMCRILRLINAMKIPSIDTHNPIFFFTQCDTNDTLDYEDPRLIQPCVC
RKIKGLRISIPLPNGTIAIAGEPRLKQSLKILNMWCLDQSLQILNSITPPGCPDG
RKHGICIRIGYKTDMDKGMHGLSTPIFYDEGYSRCSCTFIOMQDLNLLH
GTTKAPTEBELIYNTDGCPSKENSFMDLSRQOLMNVNATLTIKTYNTHGPNNA
ATSDSLISPIRLIGWPIIKTOLITTYETRILOQSOVTFEIRIDKNIQIKKTORG
VSDPSCLNRHGRNDVLYIQLKADGNLILINSTCATKCSNSKVSCHSIALE
QIKN"
/gene="U44orf"
/feature="U44orf"
/codon_start=1
/protein_id="CAA63170.1"
/db_xref="GI:1044874"
/db_xref="Swiss-Prot:P52473"
/translation="MNSLGYGVGNCRRLSLFGOPTQAKYEMLCSSDLETELTF
LDMTYKDFGVQNDIISHQDTEMTKLILGLIMYKTKLRLHVMRCLSMCPNHVAD
ALCYELMKARELIQTMVPMKTLIGFSGCTENLNOLNLFKATDOSTLSVDKINS
IEIDGENSKRLVELDPIHLEGTGVALNVVTEAPSEKVSIRVSEBES"
/complement(10155..10601)
/gene="U45orf"
/feature="U45orf"
/codon_start=1
/protein_id="CAA63171.1"
/db_xref="GI:1044875"
/db_xref="GOA:Q06095"
/db_xref="Swiss-Prot:Q06095"
/translation="EFSINFLPCYIEPPEWMINLSPPEAFILKASREFIAKNSYT
IOYDAMVCADEKLALMIPSKIAKGLIERIYMNKOTIIPSIKIFNSRKYIYPT
GICIRILIFGHCFLSLMERAINRLQVLDANSSFLVHYAASNA"
ORIGIN
Query Match 84.0%; Score 16.6; DB 14; Length 10601;
Best Local Similarity 90.0%; Pred. No. 3.5e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 AGTGGCAACAGCTACATGG 20
Db 9449 AGTCGCAACACTACGATGG 9468
RESULT 16
ACI04588 56884 bp DNA linear HTG 13-DEC-2001
LOCUS Homo sapiens chromosome 17 clone RP11-86B16 map 17, LOW-PASS
DEFINITION SEQUENCE SAMPLING.
ACCESSION ACI04588
VERSION ACI04588.1 GI:17646988
KEYWORDS HTG; HTGS PHASE0
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 56884)
AUTHORS Birren,B., Linton,L., Nusbaum,C. and Lander,E.
```

TITLE Homo sapiens chromosome 17, clone RP11-88B16
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 56884)
AUTHORS Birren, B., Linton, L., Nusbaum, C., Lander, E., Allen, A., Allen, N.,
Anderson, S., Barna, N., Bastien, V., Boguski, M., Bouckge, B.,
Brown, A., Camarata, J., Campilano, A., Chang, J., Chazotte, B.,
Choquet, Y., Colangelo, M., Collins, S., Collymore, A., Cook, A.,
Cooper, P., Dearlano, K., Dewar, K., Diaz, J., Dodge, S., Fero, S.,
Ferrera, P., Fitzhugh, W., Gage, D., Galagan, J., Gardy, S.,
Gardes, B., Gelfand, S., Geyette, M., Graham, L., Grand-Pierre, N.,
Hagos, B., Heath, A., Horton, L., Hulme, M., Iliev, I., Johnson, R.,
Jones, C., Kamat, A., Karas, A., Kells, C., Lacroque, K.,
Lamares, R., Landers, T., Lehotzky, J., Levine, R., Liu, G.,
Maclean, C., Macdonald, P., Major, J., Marquis, N., Matthews, C.,
McCarthy, M., McEwan, P., McKernan, K., McPherson, R., Melnick, D.,
Menus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C.,
Norbu, C., Norman, C. H., O'Connor, T., O'Donnell, P., O'Neill, D.,
Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Polara, V.,
Raymond, C., Retter, R., Rieback, M., Riley, R., Rise, C., Rogov, P.,
Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupbach, R.,
Seaman, S., Severy, P., Spencer, B., Stange, Thoman, N., Stojanovic, N.,
Strass, N., Subramanian, A., Talmas, J., Tesfaye, S., Theodore, J.,
Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H.,
Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W., Young, G.,
Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission
Submitted (13-DEC-2001) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: <http://www-seq.wi.mit.edu>
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: 122861
Center clone name: 88_B_16

* NOTE: This record contains 69 individual
* sequencing reads that have not been assembled into
* contigs. Runs of N are used to separate the reads
* and the order in which they appear is completely
* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene-rich and allows
* overlap relationships among clones to be deduced.
* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.

1 753: contig of 753 bp in length
* 754 853: gap of 100 bp
* 854 1596: contig of 743 bp in length
* 1597 1696: gap of 100 bp
* 1697 2438: contig of 742 bp in length
* 2439 2538: gap of 100 bp
* 2539 3278: contig of 740 bp in length
* 3279 3378: gap of 100 bp
* 3379 4106: contig of 728 bp in length
* 4107 4206: gap of 100 bp
* 4207 4890: contig of 684 bp in length
* 4891 4990: gap of 100 bp
* 4991 5701: contig of 711 bp in length
* 5702 5801: gap of 100 bp
* 5802 6548: contig of 748 bp in length
* 6549 6650: gap of 100 bp
* 6650 7353: contig of 704 bp in length
* 7354 7453: gap of 100 bp
* 7454 8144: contig of 691 bp in length
* 8145 8244: gap of 100 bp
* 8245 8991: contig of 747 bp in length
* 8992 9091: gap of 100 bp

9092 9838: contig of 747 bp in length
* 9839 9938: gap of 100 bp
* 9939 10682: contig of 744 bp in length
* 10683 10782: gap of 100 bp
* 10783 11521: contig of 739 bp in length
* 11522 11621: gap of 100 bp
* 11622 12345: contig of 724 bp in length
* 12346 14445: gap of 100 bp
* 14446 13170: contig of 725 bp in length
* 13171 13270: gap of 100 bp
* 13271 13983: contig of 713 bp in length
* 13984 14083: gap of 100 bp
* 14084 14787: contig of 704 bp in length
* 14788 14887: gap of 100 bp
* 14888 15630: contig of 743 bp in length
* 15631 15730: gap of 100 bp
* 15731 16475: contig of 745 bp in length
* 16476 16575: gap of 100 bp
* 16576 17301: contig of 726 bp in length
* 17302 17401: gap of 100 bp
* 17402 18124: contig of 723 bp in length
* 18125 18224: gap of 100 bp
* 18225 18937: contig of 713 bp in length
* 18938 19037: gap of 100 bp
* 19038 19778: contig of 741 bp in length
* 19779 19878: gap of 100 bp
* 19879 20589: contig of 711 bp in length
* 20590 20689: gap of 100 bp
* 20690 21418: contig of 729 bp in length
* 21419 21518: gap of 100 bp
* 21519 22220: contig of 702 bp in length
* 22221 22320: gap of 100 bp
* 22321 23018: contig of 698 bp in length
* 23019 23118: gap of 100 bp
* 23119 23806: contig of 688 bp in length
* 23807 23906: gap of 100 bp
* 23907 24649: contig of 743 bp in length
* 24650 24749: gap of 100 bp
* 24750 25479: contig of 730 bp in length
* 25480 25579: gap of 100 bp
* 25580 26311: contig of 732 bp in length
* 26312 26411: gap of 100 bp
* 26412 27138: contig of 727 bp in length
* 27139 27238: gap of 100 bp
* 27239 27944: contig of 706 bp in length
* 27945 28044: gap of 100 bp
* 28045 28755: contig of 711 bp in length
* 28756 28855: gap of 100 bp
* 28856 29587: contig of 732 bp in length
* 29588 29687: gap of 100 bp
* 29688 30398: contig of 711 bp in length
* 30399 30498: gap of 100 bp
* 30499 31189: contig of 691 bp in length
* 31190 31289: gap of 100 bp
* 31290 32188: contig of 899 bp in length
* 32189 32288: gap of 100 bp
* 32289 33022: contig of 734 bp in length
* 33023 33122: gap of 100 bp
* 33123 33854: contig of 732 bp in length
* 33855 33954: gap of 100 bp
* 33955 34682: contig of 728 bp in length
* 34683 34782: gap of 100 bp
* 34783 35516: contig of 724 bp in length
* 35517 35616: gap of 100 bp
* 35617 36329: contig of 713 bp in length
* 36330 36429: gap of 100 bp
* 36430 37147: contig of 718 bp in length
* 37148 37247: gap of 100 bp
* 37248 37976: contig of 729 bp in length
* 37977 38076: gap of 100 bp
* 38077 38780: contig of 704 bp in length
* 38781 38880: gap of 100 bp
* 38881 39572: contig of 692 bp in length


```
* 12450 12549: gap of 100 bp
* 12550 13302: contig of 753 bp in length
* 13303 13402: gap of 100 bp
* 13403 14163: contig of 761 bp in length
* 14164 14263: gap of 100 bp
* 14264 15035: contig of 772 bp in length
* 15036 15135: gap of 100 bp
* 15136 15905: contig of 770 bp in length
* 15906 16005: gap of 100 bp
* 16006 16769: contig of 764 bp in length
* 16770 17592: contig of 723 bp in length
* 17593 17692: gap of 100 bp
* 17693 18415: contig of 723 bp in length
* 18416 18515: gap of 100 bp
* 18516 19263: contig of 748 bp in length
* 19264 19363: gap of 100 bp
* 19364 20132: contig of 769 bp in length
* 20133 20232: gap of 100 bp
* 20233 20965: contig of 733 bp in length
* 20966 21065: gap of 100 bp
* 21066 21792: contig of 727 bp in length
* 21793 21892: gap of 100 bp
* 21893 22603: contig of 711 bp in length
* 22604 22703: gap of 100 bp
* 22704 23447: contig of 744 bp in length
* 23448 23547: gap of 100 bp
* 23549 24290: contig of 743 bp in length
* 24291 24390: gap of 100 bp
* 24391 25162: contig of 772 bp in length
* 25163 25262: gap of 100 bp
* 25263 26044: contig of 782 bp in length
* 26045 26144: gap of 100 bp
* 26145 26921: contig of 777 bp in length
* 26922 27021: gap of 100 bp
* 27022 27737: contig of 716 bp in length
* 27738 27837: gap of 100 bp
* 27838 28597: contig of 760 bp in length
* 28598 28697: gap of 100 bp
* 28699 29448: contig of 751 bp in length
* 29449 29548: gap of 100 bp
* 29549 30318: contig of 770 bp in length
* 30319 30418: gap of 100 bp
* 30419 31169: contig of 751 bp in length
* 31170 31269: gap of 100 bp
* 31270 32005: contig of 736 bp in length
* 32006 32105: gap of 100 bp
* 32106 32854: contig of 749 bp in length
* 32855 32954: gap of 100 bp
* 32955 33714: contig of 760 bp in length
* 33715 33814: gap of 100 bp
* 33815 34591: contig of 777 bp in length
* 34592 34691: gap of 100 bp
* 34692 35448: contig of 757 bp in length
* 35449 35548: gap of 100 bp
* 35549 36315: contig of 768 bp in length
* 36316 36415: gap of 100 bp
* 36416 37139: contig of 723 bp in length
* 37140 37239: gap of 100 bp
* 37240 37996: contig of 757 bp in length
* 37997 38096: gap of 100 bp
* 38097 38839: contig of 743 bp in length
* 38840 38939: gap of 100 bp
* 38940 39714: contig of 775 bp in length
* 39715 39814: gap of 100 bp
* 39815 40483: contig of 669 bp in length
* 40484 40583: gap of 100 bp
* 40584 41307: contig of 724 bp in length
* 41308 41407: gap of 100 bp
* 41408 42110: contig of 703 bp in length
* 42111 42947: gap of 100 bp
* 42211 43047: contig of 737 bp in length
* 42948 43047: gap of 100 bp
```

```
* 43048 43811: contig of 764 bp in length
* 43812 43912: gap of 100 bp
* 43912 44673: contig of 762 bp in length
* 44674 44773: gap of 100 bp
* 44774 45539: contig of 766 bp in length
* 45540 45639: gap of 100 bp
* 45640 46410: contig of 771 bp in length
* 46411 46510: gap of 100 bp
* 46511 47236: contig of 726 bp in length
* 47237 47336: gap of 100 bp
* 47337 48096: contig of 760 bp in length
* 48097 48196: gap of 100 bp
* 48197 48957: contig of 761 bp in length
* 48958 49057: gap of 100 bp
* 49058 49802: contig of 745 bp in length
* 49803 49902: gap of 100 bp
* 49903 50632: contig of 730 bp in length
* 50633 50732: gap of 100 bp
* 50733 51463: contig of 731 bp in length
* 51464 51563: gap of 100 bp
* 51564 52288: contig of 725 bp in length
* 52289 52388: gap of 100 bp
* 52389 53157: contig of 769 bp in length
* 53158 53257: gap of 100 bp
* 53258 54035: contig of 778 bp in length
* 54036 54135: gap of 100 bp
* 54136 54894: contig of 759 bp in length
* 54895 54994: gap of 100 bp
* 54995 55750: contig of 756 bp in length
* 55751 55850: gap of 100 bp
* 55851 56622: contig of 772 bp in length
* 56623 56722: gap of 100 bp
* 56723 57456: contig of 734 bp in length
* 57457 57556: gap of 100 bp
* 57557 58258: contig of 702 bp in length
```

Query Match 84.0%; Score 16.8; DB 2; Length 71049;
Best Local Similarity 90.0%; Pred. No. 3.2e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AGTCGCAACGATACATCG 20
DB 23922 AGTCACAAACGATACACGG 23941

RESULT 18
AC112029 2/c
WPCOMMENT
Sequence Split into 6 fragments LOCUS AC112029 Accession AC112029

Fragment Name Begin End
AC112029_0 1 110000
AC112029_1 1 210000
AC112029_2 200001 310000
AC112029_3 300001 410000
AC112029_4 400001 510000
AC112029_5 500001 569567
Continuation (3 of 6) of AC112029 from base 200001 (AC112029 Ratrus norvegicus clone CH2;

Query Match 84.0%; Score 16.8; DB 2; Length 110000;
Best Local Similarity 90.0%; Pred. No. 3.2e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AGTCGCAACGATACATCG 20
DB 102858 AGCAGCAACGATACATCG 102839

RESULT 19
AC112029 3/c
WPCOMMENT
Sequence Split into 6 fragments LOCUS AC112029 Accession AC112029
Fragment Name Begin End
AC112029_0 1 110000

AC112029.1 100001 210000
AC112029.2 200001 310000
AC112029.3 300001 410000
AC112029.4 400001 510000
AC112029.5 500001 569567
Continuation (4 of 6) of AC112029 from base 300001 (AC112029 Rattus norvegicus clone CH2)

Query Match 84.0%; Score 16.8; DB 2; Length 110000;
Best Local Similarity 90.0%; Pred. No. 3.2e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 AGTGGCAACGTCATCG 20
Db 2858 AGCAGCAACGTCATCG 2839

RESULT 20
OSJN00288/c 112404 bp DNA linear PLN 10-FEB-2004
LOCUS Oryza sativa genomic DNA, chromosome 4, BAC clone: OJ000315_02,
DEFINITION complete sequence.

ACCESSION AL731639
VERSION AL731639.3 GI:38569166
KEYWORDS HTG.
SOURCE Oryza sativa (japonica cultivar-group)
ORGANISM Oryza sativa (japonica cultivar-group)
Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Eriatoidae; Oryzae; Oryza.

REFERENCE

1 Feng, Q., Zhang, Y., Hao, P., Wang, S., Fu, G., Huang, Y., Li, Y., Zhu, J.,
Liu, Y., Hu, X., Jia, P., Zhang, Y., Zhao, Q., Yang, K., Yu, S., Tang, Y.,
Weng, Q., Zhang, L., Lu, Y., Mu, J., Lu, Y., Zhang, L., S., Yu, Z., Fan, D.,
Liu, X., Lu, Y., Li, C., Wu, Y., Sun, T., Lei, H., Li, F., Hu, H., Guan, J.,
Wu, M., Zhang, R., Zhou, B., Chen, Z., Chen, L., Jin, Z., Wang, R.,
Yin, H., Cai, Z., Ren, S., Lv, G., Gu, W., Zhu, G., Tu, Y., Jia, J.,
Zhang, Y., Chen, J., Kang, H., Chen, X., Shao, C., Sun, Y., Hu, Q.,
Zhang, X., Zhang, W., Wang, L., Ding, C., Sheng, H., Gu, J., Chen, S.,
Ni, L., Zhu, F., Chen, W., Lan, L., Lai, Y., Cheng, Z., Gu, M., Jiang, J.,
Li, J., Hong, G., Xue, Y., and Han, B.
Sequence and analysis of rice chromosome 4
Nature 420 (6913), 316-320 (2002)
2 12447439

AUTHORS

Fu, G., Wang, S., Ren, S., X., Lv, G., Lin, W., Gu, W., Q., Zhu, G., F.,
Tu, Y., Jia, P., Yin, H., F., Zhang, Y., Cai, Z., Chen, J., Kang, H.,
Chen, X., Y., Shao, Y., Sun, Y., Hu, Q., P., Zhang, X., L., Zhang, W.,
Weng, L., U., Ding, C., M., Sheng, H., H., Gu, J., L., Chen, S., T., Ni, L.,
Zhang, F., H., Han, B., Feng, Q., Huang, Y., C., Li, Y., Zhu, J., J., Zhao, Q.,
Hu, X., Liu, Y., L., Mu, J., Yu, Z., Chen, L., Fan, D., L., Weng, Q., J.,
Zhang, L., Lu, Y., Q., Yu, S., L., Liu, X., H., Lu, T., T., Zhang, Y., J., Lu, Y.,
Li, C., Li, T., Zhang, Y., Hu, H., Jia, P., X., Qian, Y., M., Yang, K.,
Zhou, B., Chen, Z., H., Hao, P., Zhang, L., Wu, M., Zhang, R., Q., Guan, J., P.
and Hong, G., F.

TITLE

Submitted (04-MAY-2002) Han Bin, National Center for Gene Research,
Chinese Academy of Sciences, 500# Cao Bao Road, Shanghai 200233,
CHINA. E-mail enquiries: bhan@ncgr.ac.cn. Clone requests:
bhan@ncgr.ac.cn
Oryza sativa japonica (nipponbare) genomic DNA, chromosome 4, BAC
clone: OJ000315_02.
On Dec 1, 2003 this sequence version replaced gi:32487896.
Web site: http://www.ncgr.ac.cn

COMMENT

REMARK
Summary Statistics
Assembly program: phrap

This is a complete sequence.
Genes were identified by a combination of several methods: Gene
prediction programs including Egenes (http://www.sotberry.com/),
genscan (http://CCR-081.mt.edu/GENSCAN.html), GenemarkHMM
(http://genemark.biology.gatech.edu/Genemark/), tRNAscan-SE (Sean
Beddy, http://genome.wustl.edu/eddy/tRNAscan-SE/), searches of the

FEATURES

source

complete sequence against NCBI non-redundant protein database (nr)
(ftp://ncbi.nlm.nih.gov/blast/db) and the EST database at NCGR.
Location/Qualifiers

1. 112404

/organism="Oryza sativa (japonica cultivar-group)"

/mol_type="genomic DNA"

/cultivar="Nipponbare"

/sub_species="japonica"

/db_xref="taxon:39947"

/chromosome="4"

/clone="OJ000315_02"

/clone_lib="NonIntro-OUT"

/complement(27..1985)

/gene="OJ000315_02.1"

/complement(join(27..154,455..719,816..1985))

/gene="OJ000315_02.1"

/codon_start=1

/protein_id="CAE05356.1"

/db_xref="GI:32487897"

/db_xref="TrEMBL:Q7X618"

/translation="WDDELFDKSSNDHSISSEEDMLVRSYNLNVSGYHCNIOC
PSLDIDEVDISPNKRLTNTMTSONSFCLSGAALISANTLANNICKGLIGEIL
PELDSFNSFRKIVSSPMSRLDLSTQSPVETESIFESIKMWRSSAPTVSSNF
LSTSEIKMAGAGEDRVOACSEKRMGLICGIVDNGRDAFLAVLYNDIVYL
YLECRIKOENGLGSPESLNGVSKSLTAMEFAEMEDVKEFETPRAGVLCCLTAV
EOAENDFLCWEOBMDRDPDVSVGVTVVILHGTIDICLINGDSRAVAVSSGM
DKRAVYLTEHSLLENPLEYQKLADHPNRPVYMGKIKGKATRAVGVYLRQK
LNDALWILRRNLCSPPVYVNTPHVSHVTDVDFVVGSGGLDFPESNDVVOIG
YQFMHNPIDGPAKYLIEQLLKAAKKAALTAELMRIPVGRKKYHDDVTIIVILG
NAQRTMTASTSL"

/complement(12897..15028)

/gene="OJ000315_02.2"

/complement(join(12897..13094,13461..13545,13645..13730,
14707..14731,14955..15028))

/gene="OJ000315_02.2"

/codon_start=1

/protein_id="CAE05357.3"

/db_xref="GI:38569167"

/db_xref="TrEMBL:Q7X8V5"

/translation="MTDDETFGSGYSCPIVACHVYNTYRIRHSQWSDYCLVLD
QRCDSFVYFNKVDSPRRNLRDLGTFRANKALSDPEVTKDHENSRRHGRLHV
SMVHLQRLRRVQPMPLRLRLRYAFISAQEWLLDLDKEGMKMKCTRADKIS"

17220..17643

/gene="OJ000315_02.3"

/join(17220..17290,17541..17643)

/gene="OJ000315_02.3"

/codon_start=1

/protein_id="CAE05358.3"

/db_xref="GI:38569168"

/db_xref="TrEMBL:Q7X699"

/translation="MASSSSSSMOALSIMSCTVIYKTFSPCEPVKPKMKVYKDRRAK
LPVSPKFTARVN"

/complement(24736..27684)

/gene="OJ000315_02.4"

/complement(join(24736..25083,26698..26751,27465..27561,
27556..27684))

/gene="OJ000315_02.4"

/codon_start=1

/protein_id="CAE05359.3"

/db_xref="GI:38569169"

/db_xref="TrEMBL:Q7X7P7"

/translation="WAVEFGEYSIMELKELPFRFDKXKXENSETLFMEDSWGSK
EGGRGMWGAQLRQNYRKQGVMAASNYVDIAGEBGRFGHGHSSTPTGAASPK
MRSSWSASAGSGGGGSGPKVCAPATHAGSFRCRLRRSSHGHPSSPTAAAP
AAAVQSSSRTVAQ"

/complement(28215..31978)

/gene="OJ000315_02.5"

/complement(join(28215..28502,28743..28883,29291..30228,
30405..30644,30781..31978))

/gene="OJ000315_02.5"

/codon_start=1

/protein_id="CAE05360.3"

/db_xref="GI:38569170"

/db_xref="rEMBL:07X787"
 /translation="MCCPRSGCVTIOSKKGIVCKEHNPPLAIVTIMVGDLTGQALAE
 GPTQHFKNKNTWKAKMOEKGLFHRPRLPANKLDELAEFDSIRKNGTSVHKVKKMOE
 AEAVRLEHWVTVGEVDEKVDYDPLHEVGSNIGPVYEDVMALTKTDVYRIRVGNM
 TLKSLPLMTLMTPEKLIVYKAHFKLEQJVELGMRDQAEKRVVEVEQNEPGLNFM
 QNRKKPRKEEASNNRVLKAGIKKKAIVESDENSESAQELVKLQMBIDRMALR
 VOLNEHFVKYQIQLDKIDVAGMIIEGQIREEAQEBENEKTKYQEBQMLDKSGVQRV
 ENDSVEYVDEDESPVQVCGEENNESQESDPAVAAGVLSQTNEDMTEKRSKSLR
 LLEKXKCCALAEADTNLTIKNLEIAINLIFKEMNTSQKSGQVNGENLSININ
 IDEVHSDCMSPDYAENMTILQSNSSEKQDPFGIOETVQNDPFAKXLDLIGGGG
 KPSNMTIPKESGCMLVGTEDEKMEIIEVSKKNPLFKAKRNDKFPWRLITVYA
 AQPEKKEEFLBLADQCYGDEPPLIGDVLIDQENKIKRGGVNMKSJFNSITEN
 ANLRMALEGGKFFPANNKQSVTEKLDRLILFNDKMEIKPEPLACGRVLERYISDHTV
 LINVGEKMLSLKPFENSMWMBEDININVAOVNSQINSSIDKXOMIQKRRKL
 KGMNANMGMYKRGKHEIMEKLEIDKCEVAGMTIIEKRGGLPEELKKNVDMPLK
 ITDFPKRIPIVEKRLNGVITLIPKCKEARMIOKRRPCLNAKQFPIDCWMILKVKV
 GKKVAIRNDDQHYETHTHKGLRQSDPLPLFNIAADALITLKRKEGGLIKGJGM
 DYMRGVAILDQAHHTICLI"
 3650. .36766
 /gene="OJ000315_02.6"
 join(36050. .36423,36754. .36766)
 /gene="OJ000315_02.6"
 /codon_start=1
 /protein_id="CAE05363.3"
 /db_xref="GI:38569171"
 /db_xref="rEMBL:07X788"
 /translation="MERARKSSSGGTGLGGAHNGARQDQGSANATGCAVRNMG
 ATGMRRARRGGTAVAIVAGARRMTARRDNDQDGAGLAARCGDGDGREAEAR
 RAGRGNAKEGLIGDVEHRRYS"
 complement(40401. .42688)
 /gene="OJ000315_02.7"
 complement(join(40401. .40769,42056. .42688))
 /gene="OJ000315_02.7"
 /codon_start=1
 /protein_id="CAE05362.3"
 /db_xref="GI:38569172"
 /db_xref="rEMBL:07X789"
 /translation="MSTSHSSLSLILSLPSSILPLGPTTRVVGAGELTMRMR
 AMSAPLDRSLTEAKVSRBEGLTEIASGPMGGRSLMAYGSEVPRPKSGREYK
 ALPSARVREIRIPAREGTPASAPFPCSAVSEVREHSGSVPASTILLARSSFT
 DPDRPSQYTAKEPHLTSHPHACGIDTLRSMKASSENKSTSAAMMRRTTLGACFL
 EURLASWVPIWSAKTKPSRREBAVDLVWATRFGRSRSLSAAMMRRTTLGACFL
 RGRGGEPPENCDDGMVEGEGVATKEVVRVAKELRGSDRAADEVAEEBARVC"
 complement(43876. .45030)
 /gene="OJ000315_02.8"
 complement(join(43876. .44247,44336. .44561,44657. .45030))
 /gene="OJ000315_02.8"
 /codon_start=1
 /protein_id="CAE05363.3"
 /db_xref="GI:38569173"
 /db_xref="rEMBL:07X653"
 /translation="VAAVPTFAVVAALVALCGGGGAGADMNENARNRGFFYGQIG
 GGVGGLIPQSDVLELFPNLNLEYAEFFCMSALGYLDGIDLSLGGGAPVGAQDQ
 ALAPFVBDIAPOFCYQEVGHLEIKQNVAGRPRLDILSANGKIYEIMTMDLPP
 PNPNPENGINPLASVYIIPVNGITGVCANRPLILPQAKIIVAGLIGESQADVITAA
 LLYHGLSRVASYGVGVAELTHISELNVJGRKVRQBGIVVAPGQSPBQQTGNII
 AADRSLAIDRTPEEILGVVIGSGDPARAGGFFPGADGRITARAFA"
 complement(46275. .46673)
 /gene="OJ000315_02.9"
 complement(46275. .46673)
 /gene="OJ000315_02.9"
 /codon_start=1
 /protein_id="CAE05364.3"
 /db_xref="GI:38569174"
 /db_xref="rEMBL:07X705"
 /translation="MSEHAVKEDVVRRLKLSAKGACVAINAALLEISFPLALKC
 KKEEASLARRARVPEKIGTSGVLAKERLVRTRTRVALVTRPTEKTIIDTRKRLN
 VEHRIPNKKKLGDRDNEGAADVNRDVCV"
 51378. .51761
 /gene="OJ000315_02.10"
 51378. .51761
 /gene="OJ000315_02.10"
 /gene="OJ000315_02.10"

Query Match	84.0%	Score 16.8	DB 8	Length 112404
Best Local Similarity	90.0%	Pred. No. 3.1e+02		
Matches	18	Conservative	0	Mismatches 2; Indels 0; Gaps 0;

QY

1

AGTCGAAACAGTACAAATGG 20

Db

11288

AGTACGAAACAGTACAAATGG 11269

RESULT 21	AL139153/c	115585 bp	DNA	linear	HTG 10-UTL-2001
LOCUS	AL139153				
DEFINITION	Homio sapiens chromosome 1 clone RP4-753A11, 10 unordered pieces.				
ACCESSION	AL139153				
VERSION	AL139153.4 GI:9863525				
KEYWORDS	HTG; HTGS; PHASE1; HTGS_CANCELLED.				
SOURCE	Homio sapiens (human)				
ORGANISM	Homio sapiens				
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homio.				
REFERENCE	1				
AUTHORS	Plumb, B.				
TITLE	Direct Submission				
JOURNAL	Submitted (03-JUL-2001) Sanger Centre, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk				
	requests: clonerequests@sanger.ac.uk				
	On Aug 21, 2000 this sequence version replaced gi:9212193.				
COMMENT	----- Genome Center				
	Center: Sanger Centre				
	Center code: SC				
	Web site: http://www.sanger.ac.uk				
	Contact: humquerry@sanger.ac.uk				
	----- Project Information				
	Center project name: dt753A11				
	----- Summary Statistics				
	Assembly program: XGAP; version 4.5				
	Sequencing vector: plasmid; L08152; 100% of reads				
	Chemistry: Dye-terminator Big Dye; 100% of reads				
	Consensus quality: 110424 bases at least Q40				
	Consensus quality: 112822 bases at least Q30				
	Consensus quality: 114272 bases at least Q20				
	Insert size: 115685; sum-of-ctnigs				
	Insert size: 127770; 7.3% error; agarose-fp				
	Quality coverage: 3.64x in Q20 bases; sum-of-ctnigs Quality				
	coverage: 3.37x in Q20 bases; agarose-fp				

	* NOTE: This is a 'working draft' sequence. It currently				
	* consists of 10 ctnigs. The true order of the pieces				
	* is not known and their order in this sequence record is				
	* arbitrary. Gaps between the ctnigs are represented as				
	* runs of N, but the exact sizes of the gaps are unknown.				
	* This record will be updated with the finished sequence				
	* as soon as it is available and the accession number will				
	* be preserved.				
	1	13312; ctnig of 13212 bp in length			
	13213	gap of 100 bp			
	13313	20141; ctnig of 6829 bp in length			
	20142	gap of 100 bp			
	20242	29470; ctnig of 9229 bp in length			
	29471	gap of 100 bp			
	29571	35119; ctnig of 5949 bp in length			
	35520	gap of 100 bp			
	35620	52338; ctnig of 16719 bp in length			
	52339	gap of 100 bp			
	52439	88864; ctnig of 36526 bp in length			
	88965	gap of 100 bp			
	89065	97388; ctnig of 8324 bp in length			
	97389	gap of 100 bp			
	97489	108993; ctnig of 11505 bp in length			
	108994	gap of 100 bp			
	109094	113851; ctnig of 4858 bp in length			
	113952	gap of 100 bp			
	114051	gap of 100 bp			
	114052	116585; ctnig of 2534 bp in length.			

```

FEATURES
Source
1. .116585
/location/Qualifiers
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="1"
/clone="RP4-753A11"
/clone_lib="RPCI-4"
1. .13212
/misc_feature
/note="assembly_fragment:01103
fragment_chain:1"
1313. .20141
/misc_feature
/note="assembly_fragment:00150
fragment_chain:1"
20242. .29470
/misc_feature
/note="assembly_fragment:00511.0"
29571. .35519
/misc_feature
/note="assembly_fragment:00511.1"
35620. .52358
/misc_feature
/note="assembly_fragment:00628"
52439. .88964
/misc_feature
/note="assembly_fragment:00071
fragment_chain:2"
89065. .97388
/misc_feature
/note="assembly_fragment:00168
fragment_chain:2"
97489. .108993
/misc_feature
/note="assembly_fragment:00346
fragment_chain:2"
109094. .119951
/misc_feature
/note="assembly_fragment:00939
fragment_chain:2"
114052. .116585
/misc_feature
/note="assembly_fragment:00581
fragment_chain:2
clone_end:SP6
vector_side:right"

ORIGIN
Query Match 84.0%; Score 16.8; DB 2; Length 116585;
Best Local Similarity 90.0%; Pzed.No. 3.1e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 AGTCGCAACACGACACATGC 20
Db 52779 AATCGACAAACGACACATGC 52760

RESULT 22
AC006945 120311 bp DNA linear ROD 17-SNP-2003
LOCUS AC006945
DEFINITION Mus musculus clone ct7-453113 map 6 strain 129/Sv, complete
sequence.
AC006945
VERSION AC006945.17 GI:34787443
KEYWORDS HTG.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 120311)
Chen,F., Do,T., Do,A., Mcdermid,H. and Roe,B.A.
Mus musculus Chromosome 6 BAC Clone ct7-453113 In CES Region
Unpublished
2 (bases 1 to 120311)
Chen,F., Do,T., Do,A., Mcdermid,H. and Roe,B.A.
Direct Submission
Submitted (02-MAR-1999) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
3 (bases 1 to 120311)
Chen,F., Do,T., Do,A., Mcdermid,H. and Roe,B.A.
Direct Submission
TITLE
AUTHORS
JOURNAL
REFERENCE

```

```

JOURNAL
Submitted (07-JUL-1999) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
REFERENCE
AUTHORS
TITLE
JOURNAL
Direct Submission
Submitted (24-JUL-1999) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
5 (bases 1 to 120311)
Chen,F., Do,T., Do,A., McDermid,H. and Roe,B.A.
Direct Submission
Submitted (15-FEB-2000) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
6 (bases 1 to 120311)
Chen,F., Do,T., Do,A., McDermid,H. and Roe,B.A.
Direct Submission
Submitted (22-FEB-2000) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
7 (bases 1 to 120311)
Chen,F., Do,T., Do,A., McDermid,H. and Roe,B.A.
Direct Submission
Submitted (15-JUN-2000) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
8 (bases 1 to 120311)
Chen,F., Do,T., Do,A., McDermid,H. and Roe,B.A.
Direct Submission
Submitted (18-JAN-2002) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
9 (bases 1 to 120311)
Chen,F., Do,T., Do,A., McDermid,H. and Roe,B.A.
Direct Submission
Submitted (25-AUG-2002) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
10 (bases 1 to 120311)
Chen,F., Do,T., Do,A., McDermid,H. and Roe,B.A.
Direct Submission
Submitted (19-OCT-2002) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
11 (bases 1 to 120311)
Chen,F., Do,T., Do,A., McDermid,H. and Roe,B.A.
Direct Submission
Submitted (17-SEP-2003) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
On Sep 17, 2003 this sequence version replaced gi:24137451.
----- Genome Center
Center: Department Of Chemistry And Biochemistry
The University Of Oklahoma
Center code:UOKNOR
-----
FEATURES
source
1..120311
Location/Qualifiers
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="129/Sv"
/db_xref="taxon:10090"
/map="6"
/clone="cct7-453113"
/clone_11b="ClbDC7 mouse BAC library"
ORIGIN
Query Match 84.0%; Score 16.8; DB 10; Length 120311;
Best Local Similarity 90.0%; Pred. No. 3.1e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0.
1 AGTCGCAACAGTACATGG 20

```



```
repeat_region      /rpt_family="tandem"
                    13953..13980
repeat_region      /rpt_family="AT-rich"
                    14964..15017
repeat_region      /rpt_family="CT-rich"
                    complement(15388..15741)
gene               /gene="OU112_G08.4"
                    /note="predicted by fgenesh"
                    complement(<15388..>15741)
mRNA               /gene="OU112_G08.4"
                    complement(15388..15741)
CDS                /gene="OU112_G08.4"
                    /product="hypothetical protein"
                    /protein_id="AAP68357.1"
                    /db_xref="GI:31712051"
                    /translation="MAEDVAGGVVRLVAABRAEDRGVAAEAGARQGRTRGGCG
                    EAYGGVGGGGELVEGARREGIGIGDSGEAREVSVDDGGGGIGRRSPFEQR
                    CGSKGRPTAAAG"
repeat_region      15570..15699
                    /rpt_family="(CCG)n"
repeat_region      15853..15892
                    /rpt_family="tandem"
repeat_region      15903..15951
                    /rpt_family="GA-rich"
repeat_region      17980..18000
                    /rpt_family="AT-rich"
repeat_region      19579..19608
                    /rpt_family="AT-rich"
repeat_region      20714..20865
                    /rpt_family="(CCTCG)n"
repeat_region      20920..20965
                    /rpt_family="GA-rich"
gene               20961..25995
                    /gene="OU112_G08.5"
                    /note="nearly identical to phytochrome C GB:BAU74448
                    GI:4190974 (Plant Physiol 118, 1535 (1998)) (Oryza sativa
                    (japonica cultivar group)); EST CB626698, CB668936,
                    CB662420 from this gene"
                    join(20961..23324,23429..24245,24639..24932,25491..25995)
mRNA               join(21260..23324,23429..24245,24639..24932,25491..25728)
CDS                /gene="OU112_G08.5"
                    /codon_start=1
                    /product="phytochrome C"
                    /protein_id="AAP68360.1"
                    /db_xref="GI:31712054"
                    /translation="MSSSRNNRATCSRSSRSKRKSARVAQTMDQIARFEGSQ
                    RHPDYSSVGAANRSGATTSNVAIQQNQRFVQFSCILAVHPETPALAIYENR
                    AEMLDLTPHVPITDQREALAVGTIVRTLFRSHFVALQKATFGDVNLNLIIVHAR
                    TSGKPPYAIWHRIDVGLVIDLEFVNPVDLPVATGAIKSKYLAAPAIARLQELPSGNL
                    SLICDYLVRVSELGTGYDRMAKFFEDHGEVIAECKRSDLEPYLGAIYPATDIPQA
                    SRPLPMNKVRMTCDGSAIPVKIIDDLSLTOPISIGSTLRAPHGCHAYMAGSVA
                    SIWSTYINDEDDDDGDSQDQPKRKIMGLMVCHHTSPRPVPLRYACDFLLQVF
                    GLOINKVELAAAKERHILRTTLLCDMLDAPVIGITQSPNNMDYKCDGALLY
                    QNOLWVGSTPSEAEIKNIIVAMQBYHDSGTSDSLVBAQIPGAALGDVYCGMAA
                    IKISKQFIWFPSHAKGIKWGAHHEPIDADNKRKHPGSSFFAFLFEVKKMSVP
                    MEDVENDAIHSILQILIRGLOEDANKNNNAKSIYAPSDMKIKIGLLELTVMEN
                    VRLIETATAIIVDITGSLINGNNKAALGTGIPWEAIGKPLVDLVIDDSEVAVKOI
                    LNALGIEIEQNQIKLKTENNHEGNNPVILWVNAACSRDLSKVGVGVFAQDVTGQ
```

Query Match 84.0%; Score 16.8; DB 8; Length 128234;
Best Local Similarity 90.0%; Pred. No. 3.1e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AGTCGCAACAGTACATGG 20
Db 94243 AGTAGCAACATACATGG 94262

RESULT 24
AF377947/C AF377947 137580 bp DNA linear PLN 30-JAN-2003

DEFINITION Oryza sativa (japonica cultivar-group) chromosome 3, BAC clone
OSJNBa0032E21, complete sequence.
ACCESSION AF377947
VERSION AF377947.3 GI:21070919
KEYWORDS HTG.
SOURCE Oryza sativa (japonica cultivar-group)
ORGANISM Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehharitidae; Oryzaceae; Oryza.
REFERENCE 1 (bases 1 to 137580)
Eastman,A.P., Smith,S.C., Bertin,N., Liang,C., Najjar,F.Z.,
Pratt,L.H. and Cordomier-Pratt,M.-M.
AUTHORS Unpublished
TITLE 2 (bases 1 to 137580)
REFERENCE Eastman,A.P., Smith,S.C., Gingle,A., Pratt,L.H. and
AUTHORS Cordomier-Pratt,M.-M.
TITLE Direct Submission
JOURNAL Submitted (01-MAY-2001) Botany, University of Georgia, Miller Plant
Sciences, Athens, GA 30602, USA
3 (bases 1 to 137580)
REFERENCE Eastman,A.P., Smith,S.C., Bertin,N., Liang,C., Najjar,F.Z.,
AUTHORS Pratt,L.H. and Cordomier-Pratt,M.-M.
TITLE Direct Submission
JOURNAL Submitted (22-MAY-2002) Botany, University of Georgia, 3604 Miller
Plant Sciences, Athens, GA 30606, USA
REMARK Sequence update by submitter
COMMENT On May 22, 2002 this sequence version replaced GI:15384618.
The orientation of the sequence is from M13-21 to Jembrev of the BAC
clone. The assembly has been confirmed via restriction digest.
Genes were predicted from the integrated results of the following:
GENSCAN1.0, Genemark 2.2. The genomic sequence and predicted gene
sequences were searched against the NCBI non-redundant nucleotide
database, SWISSPROT, and NCBI EST database. Protein similarities of
the coding regions were searched against SWISSPROT with BLASTP2.0.
Genes are annotated based on the level of evidence. Genes with
high homology hits to known proteins were annotated as the name of
the protein hit. Genes with no protein data but high EST homology
were annotated as unknown proteins. Predicted proteins are genes
predicted over most of their length by both Genscan and Genemark.

FEATURES
source Location/Qualifiers
1..137580
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="genomic DNA"
/db_xref="taxon:39947"
/chromosome="3"
/clone="OSJNBa0032E21"
/clone_1kb="BAC"
1266..1420
/rpt_unit="c9g"
repeat_region 1749..1925
/rpt_unit="c9g"
repeat_region complement(2314..2374)
/rpt_unit="c9g"
repeat_region complement(2846..2902)
/rpt_unit="c9g"
LTR 5573..7073
/note="High similarity BLASTN hit to LTR of Oryza sativa
PAC clone P0416G11"
complement(9044..9230)
repeat_region /rpt_unit="c9g"
repeat_region complement(14921..15083)
/rpt_unit="c9g"
gene <16521..>18404
/gene="OSJNBa0032E21.01"
join(<16521..17973,18307..>18404)
/gene="OSJNBa0032E21.01"
/product="unknown protein"
join(16521..17973,18307..18404)
/gene="OSJNBa0032E21.01"
/note="confirmed by high quality EST hit to GB acc_num
AU0031009"


```

repeat_region      /rpt_family="L1a"
                    /rpt_family="Tigger2a"
repeat_region      /rpt_family="Tigger2a"
                    /rpt_family="TTCCLn"
repeat_region      complement(12311,.12594)
                    /rpt_family="Aiufo"
repeat_region      12596,.12632
                    /rpt_family="Tigger2a"
repeat_region      12875,.12914
                    /rpt_family="AT_rich"
repeat_region      complement(13242,.13576)
                    /rpt_family="Aiuu"
repeat_region      13780,.13861
                    /rpt_family="LIME3a"
repeat_region      complement(13888,.14298)
                    /rpt_family="MT1A1"
repeat_region      14303,.14610
                    /rpt_family="LIME3a"
repeat_region      14935,.15244
                    /rpt_family="L1P"
repeat_region      15245,.15608
                    /rpt_family="THE1B"
repeat_region      15609,.17164
                    /rpt_family="THE1B-irc"
repeat_region      17165,.17528
                    /rpt_family="THE1B"
repeat_region      17325,.18050
                    /rpt_family="L1P"
repeat_region      18028,.18163
                    /rpt_family="LIME3"
repeat_region      complement(18164,.18586)
                    /rpt_family="MERS7A"
repeat_region      18587,.18721
                    /rpt_family="LIME3"
repeat_region      18730,.18780
                    /rpt_family="(GA)n"
repeat_region      complement(18811,.19233)
                    /rpt_family="L1"
repeat_region      complement(19234,.19947)
                    /rpt_family="L1PA10"
repeat_region      complement(19948,.20072)
                    /rpt_family="L1"
repeat_region      complement(20088,.22341)
                    /rpt_family="L1"
repeat_region      complement(22361,.22521)

```

Query	Match	Similarity	Score	DE	Length
Best Local	Similarity	90.0%	Pred. No. 3.1e+02		
Matches	18; Conservative	0	Mismatches	2	Indels 0; Gaps 0

Db	1	AGTGGCAACAGTACCATG 20
13683	AGTCACAAACAGTACCAAGG 13702	

RESULT 26	LOCUS	HHV6AGNM	DEFINITION	VERSION	KEYWORDS
HHV6AGNM	159321 bp	DNA	linear	VRL 17-FEB-1997	
Human herpesvirus-6 (HHV-6) U1102, variant A DNA, complete virion genome.					
X83413.1	GI:853961				
alkaline exonuclease; complete genome; DNA polymerase; DR2; DR3; DR4; DR5; DR6; DR7; DR8; dUTPase; ganciclovir kinase; GCH; GH; GL; glycoprotein B; glycoprotein gp82/105; glycoprotein H; glycoprotein I; glycoprotein M; gM; helicase/primase complex; IE glycoprotein; IE1; integral membrane protein; large tegument protein; L1; L1r; major capsid protein; major DNA binding protein; membrane protein; origin binding protein; pBS1; phosphotransferase; pol processivity factor; pPol10; pp41; replication protein; R1; transactivator; U1; U10; U100; U11; U12; U13; U14; U15; U16; U17; U18; U19; U2; U20; U21; U22; U23; glycoprotein; U24; U25; U26; U27; U28; U29; U3; U30;					

SOURCE ORGANISM	Human herpesvirus 6 Human herpesvirus 6 varicases, dsDNA viruses, no RNA stage; Herpesviridae; Betaherpesvirinae; Roseolovirus.
REFERENCE AUTHORS TITLE	1 Chee,M.S., Lawrence,G.J., and Barrell,B.G. Alpha-, beta- and gammaherpesviruses encode a putative phosphotransferase
JOURNAL MEDLINE PUBMED	J. Gen. Virol. 70 (Pt 5), 1151-1160 (1989) 89279291 2543772
REFERENCE AUTHORS	2 Lawrence,G.L., Chee,M., Craxton,M.A., Gompels,U.A., Honess,R.W. and Barrall,B.G. Human herpesvirus 6 is closely related to human cytomegalovirus
JOURNAL MEDLINE PUBMED	J. Virol. 64 (1), 287-299 (1990) 90080132 2152817
REFERENCE AUTHORS	3 Littleler,E., Lawrence,G., Liu,M.Y., Barrall,B.G. and Arrand,J.R. Identification, cloning, and expression of the major capsid protein gene of human herpesvirus 6
JOURNAL MEDLINE PUBMED	J. Virol. 64 (2), 714-722 (1990) 90112641 2153237
REFERENCE AUTHORS	4 Martin,M.E., Thomson,B.J., Honess,R.W., Craxton,M.A., Gompels,U.A., Liu,M.Y., Littleler,E., Arrand,J.R., Teo,I. and Jones,M.D. The genome of human herpesvirus 6: maps of unit-length and concatemeric genomes for nine restriction endonucleases
JOURNAL MEDLINE PUBMED	J. Gen. Virol. 72 (Pt 1), 157-168 (1991) 1846644
REFERENCE AUTHORS	5 Thomson,B.J., Efsthaliou,S. and Honess,R.W. Acquisition of the human adeno-associated virus type-2 rep gene by human herpesvirus type 6
JOURNAL MEDLINE PUBMED	Nature 351 (6321), 78-80 (1991) 91226542 1851252
REFERENCE AUTHORS	6 Chang,C.K. and Balachandran,N. Identification, characterization, and sequence analysis of a cDNA encoding a phosphoprotein of human herpesvirus 6
JOURNAL MEDLINE PUBMED	J. Virol. 65 (6), 2884-2894 (1991) 91237802 1851860
REFERENCE AUTHORS	7 Teo,I.A., Griffin,B.E. and Jones,M.D. Characterization of the DNA polymerase gene of human herpesvirus 6
JOURNAL MEDLINE PUBMED	J. Virol. 65 (9), 4670-4680 (1991) 91333007 1651403
REFERENCE AUTHORS	8 Martin,M.E., Nicholas,J., Thomson,B.J., Newman,C. and Honess,R.W. Identification of a transactivating function mapping to the putative immediate-early locus of human herpesvirus 6
JOURNAL MEDLINE PUBMED	J. Virol. 65 (10), 5381-5390 (1991) 91374590 1654446
REFERENCE AUTHORS	9 Gompels,U.A., Cars,A.L., Sun,N. and Arrand,J.R. Infectivity determinants encoded in a conserved gene block of human herpesvirus-6
JOURNAL MEDLINE	DNA Seq. 3 (1), 25-39 (1992) 93091236

REFERENCE	PUBMED	1333836
AUTHORS		Geng Y.O., Chandran, B., Josephs, S.F. and Wood, C.
TITLE		Identification and characterization of a human herpesvirus 6 gene segment that trans activates the human immunodeficiency virus type 1 promoter
JOURNAL		J. Virol. 66 (3), 1564-1570 (1992)
MEDLINE		92148942
PUBMED		1310766
REFERENCE		11
AUTHORS		Neipel, F., Ellinger, K. and Fleckenstein, B.
TITLE		Gene for the major antigenic structural protein (p100) of human herpesvirus 6
JOURNAL		J. Virol. 66 (6), 3918-3924 (1992)
MEDLINE		92260671
PUBMED		1374813
REFERENCE		12
AUTHORS		Thomson, B.J. and Honess, R.W.
TITLE		The right end of the unique region of the genome of human herpesvirus 6 U1102 contains a candidate immediate early gene enhancer and a homologue of the human cytomegalovirus US22 gene family
JOURNAL		J. Gen. Virol. 73 (Pt 7), 1649-1660 (1992)
MEDLINE		92333248
PUBMED		1321205
REFERENCE		13
AUTHORS		Eftachidou, S., Lawrence, G.L., Brown, C.M. and Barrel, B.G.
TITLE		Identification of homologues to the human cytomegalovirus US22 gene family in human herpesvirus 6
JOURNAL		J. Gen. Virol. 73 (Pt 7), 1661-1671 (1992)
MEDLINE		92333249
PUBMED		1321206
REFERENCE		14
AUTHORS		Qian, G., Wood, C. and Chandran, B.
TITLE		Identification and characterization of glycoprotein gH of human herpesvirus-6
JOURNAL		J. Virol. 194, 380-386 (1993)
MEDLINE		92333248
PUBMED		1321205
REFERENCE		15
AUTHORS		Ellinger, K., Neipel, F., Foà-Tomasi, L., Campadelli-Fime, G. and Fleckenstein, B.
TITLE		The glycoprotein B homologue of human herpesvirus 6
JOURNAL		J. Gen. Virol. 74 (Pt 3), 495-500 (1993)
MEDLINE		93187613
PUBMED		8383182
REFERENCE		16
AUTHORS		Gompels, U.A., Carrigan, D.R., Carsz, A.L. and Arno, J.
TITLE		Two groups of human herpesvirus 6 identified by sequence analyses of laboratory strains and variants from Hodgkin's lymphoma and bone marrow transplant patients
JOURNAL		J. Gen. Virol. 74 (Pt 4), 613-622 (1993)
MEDLINE		93324882
PUBMED		8385692
REFERENCE		17
AUTHORS		Pfeiffer, B., Berneman, Z.N., Neipel, F., Chang, C.K., Tiwathapong, S. and Chandran, B.
TITLE		Identification and mapping of the gene encoding the glycoprotein complex gp82-gp105 of human herpesvirus 6 and mapping of the neutralizing epitope recognized by monoclonal antibodies
JOURNAL		J. Virol. 67 (8), 4611-4620 (1993)
MEDLINE		93523202
PUBMED		7687301
REFERENCE		18
AUTHORS		Pellet, P.E., Sanchez-Martinez, D., Dominguez, G., Black, J.B., Anton, E., Grenmeyer, C. and Dambaugh, T.R.
TITLE		A strongly immunoreactive virion protein of human herpesvirus 6 variant B strain Z29: identification and characterization of the gene and mapping of a variant-specific monoclonal antibody reactive epitope
JOURNAL		Virology 195 (2), 521-531 (1993)
MEDLINE		93331710
PUBMED		7687803
REFERENCE		19
AUTHORS		Liu, D.X., Gompels, U.A., Nicholas, J. and Leilic, C.


```

repeat_region /note="MLT2B repeat: matches 3. .404 of consensus"
42057. .42212
repeat_region /note="3 copies 52 mer 84% conserved"
42058. .42237
repeat_region /note="30 copies 6 mer tatata 78% conserved"
42114. .42233
repeat_region /note="30 copies 4 mer tata 87% conserved"
42236. .42291
misc_feature /note="MLT2B repeat: matches 394. .448 of consensus"
42053. .43470
/note="match: GSS: Em:AQ827367"
misc_feature /note="match: GSS: Em:AQ150705"
44413. .44860
repeat_region /note="Tiger2a repeat: matches 2. .429 of consensus"
43405. .43438
repeat_region /note="17 copies 2 mer tt 82% conserved"
47212. .47647
repeat_region /note="MER21B repeat: matches 339. .789 of consensus"
47695. .48392
repeat_region /note="MER50 repeat: matches 1. .722 of consensus"
48393. .48643
repeat_region /note="MER21B repeat: matches 78. .329 of consensus"
50121. .50192
repeat_region /note="36 copies 2 mer at 91% conserved"
50941. .51315
repeat_region /note="TRHC repeat: matches 1. .371 of consensus"
54119. .54164
repeat_region /note="MER83-internal repeat: matches 3088. .3134 of consensus"
54264. .54339
repeat_region /note="MER83-internal repeat: matches 2896. .3088 of consensus"
54624. .54663
repeat_region /note="20 copies 2 mer tt 77% conserved"
55288. .55355
repeat_region /note="18 copies 4 mer cata 75% conserved"
55303. .55356
repeat_region /note="9 copies 6 mer acacac 79% conserved"
56737. .57186
misc_feature /note="match: GSS: Em:AQ834443"
56794. .56943
repeat_region /note="75 copies 2 mer aa 58% conserved"
56856. .57167
misc_feature /note="match: GSS: Em:AQ694480"
57634. .57815
repeat_region /note="91 copies 2 mer aa 56% conserved"
57843. .57903
repeat_region /note="L1M4A repeat: matches 6215. .6273 of consensus"
58125. .58409
misc_feature /note="match: GSS: Em:AQ392809"
complement(58148. .58503)
misc_feature /note="match: STS: Em:G06380"
58224. .58502
/note="match: STS: Em:G11051 Em:G11052"
58404. .58429
repeat_region /note="13 copies 2 mer aa 100% conserved"
60042. .60259
repeat_region /note="L1P repeat: matches 2859. .3076 of consensus"
60261. .60765
/note="L1PAs repeat: matches 5641. .6145 of consensus"
61895. .61942
repeat_region /note="24 copies 2 mer tt 81% conserved"
61896. .61943
repeat_region /note="12 copies 4 mer ttgg 83% conserved"
68755. .69139
repeat_region /note="MSTC repeat: matches 1. .405 of consensus"
70484. .70531
repeat_region /note="24 copies 2 mer tt 81% conserved"
74889. .75634
/misc_feature /note="LTR9 repeat: matches 10. .625 of consensus"
complement(75294. .75722)
/note="match: GSS: Em:AQ039479"

```

```

repeat_region 81182. .83257
/note="L1P repeat: matches 1656. .3746 of consensus"
repeat_region 83243. .84833
/note="L1P10 repeat: matches 4086. .5694 of consensus"
repeat_region 85449. .85939
/note="L1P10 repeat: matches 5698. .6165 of consensus"
misc_feature complement(88001. .88737)
/note="match: GSS: Em:AQ253332"
complement(88277. .88711)
misc_feature /note="match: GSS: Em:AQ622463"
90455. .90542
repeat_region /note="22 copies 4 mer gtgt 77% conserved"
90485. .90544
repeat_region /note="30 copies 2 mer gt 86% conserved"
90767. .92294
/note="Cpg island"
misc_feature /evidence=not_experimental
91748. .91795
repeat_region /note="8 copies 6 mer cgcgc 81% conserved"
91751. .91804
repeat_region /note="18 copies 3 mer cgc 79% conserved"
92549. .92576
repeat_region /note="14 copies 2 mer ac 96% conserved"
93377. .93410
repeat_region /note="17 copies 2 mer aa 85% conserved"

```

```

Query Match 84.0%; Score 16.8; DB 9; Length 161040;
Best Local Similarity 90.0%; Pred. No. 3,1e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```

Qy 1 AGTCGCAACAGTACATCG 20
Db 152329 AATCACAACAGTACATCG 152348

RESULT 29
AB021506 161573 bp DNA linear VRL 24-OCT-2000
LOCUS Human herpesvirus 6 DNA, variant B, complete genome, strain:HST.
DEFINITION AB021506
ACCESSION AB021506
VERSION AB021506.1 GI:4995977
KEYWORDS Human herpesvirus 6
SOURCE Human herpesvirus 6
ORGANISM Human herpesvirus 6
Virus; dsDNA viruses, no RNA stage; Herpesviridae;
Betaherpesvirinae; Roseolovirus.
REFERENCE 1 (bases)
Isegawa,Y., Mukai,T., Nakano,K., Kagawa,M., Chen,J., Mori,Y.,
and Yamashita,K.
Sunagawa,T., Kawanishi,K., Sashihara,J., Hata,A., Zou,P., Kousuge,H.
Comparison of the complete DNA sequences of human herpesvirus 6
variants A and B
J. Virol. 73 (10), 8053-8063 (1999)
JOURNAL 99412319
MEDLINE 10482554
PUBMED 2 (bases 1 to 161573)
REFERENCE 2
AUTHORS Isegawa,Y.
TITLE Direct Submission
JOURNAL Submitted (18-DEC-1998) Yuji Isegawa, Osaka University Medical
School, Department of Microbiology; Ymada-oka 2-2, Suita, Osaka
565-0871, Japan (E-mail:iisegawa@micro.med.osaka-u.ac.jp,
Tel.81-6-879-3323, Fax:81-6-879-3329)
FEATURES
source location/Qualifiers
1.161573
/organism="Human herpesvirus 6"
/virus
/mol_type="genomic DNA"
/strain="HST"
/db_xref="taxon:10368"
/prop_variant="B"
complement(18. .365)
/gene="UL1"
complement(18. .365)
CDS

```


CDS
complement (9467, 10768)
/gene="U2"
/note="92% identical to U2 of strain U102 of HHV-6
US22 gene family"
/codon_start=1
/protein_id="BA47822.1"
/db_xref="GI:4995989"
/translation="MCSQPCGASVHFFHGRGRREYDSTLFVNSCSVLNTVIVYVR
LPTVLPISFEFFILILSRIPHRKMPCKSPPLGISWSLALSPSSCSFSGARD
LRCDAAVEVKTAFVITLVARLSADVDLFTSTHCRALSMVGAELFATSDM
LGIIOALAKLSRGYCCPGMDLTGIVTFAVSALVVGCGEITVANGVDALY
RLADAGLKMHLRREPRVYSGKCMETGASFGMGVDLAPAVSPDALVLP
KMGAPPEFVAPRRARERWRLIGGVAIVIGFPGRGVTLPLLRORVIMONQVY
AASLDGAVIRLDSFRATLAKGVKLFKXHRPPEHLTMQILPTTCHAPVINTPAV
YQLSPHVREKSAVSCGASTVVRDCEDTLRDAGVDS"
/complement (10891, 12051)
/gene="U3"
complement (10891, 12051)
CDS
Query Match 64.0%; Score 16.8; DB 14; Length 161573;
Best Local Similarity 90.0%; Pred. No. 3.1e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 AGTCGCAACAGTACATGG 20
Db 74271 AGTCGCAACAGTACATGG 74290
RESULT 30
AF157706 162114 bp DNA linear VRL 15-OCT-2003
DEFINITION Human herpesvirus 6B strain Z29, complete genome.
ACCESSION AF157706.1 L13162 L14772 L16947
VERSION AF157706.1 GI:5733510
KEYWORDS
SOURCE Human herpesvirus 6B
ORGANISM Human herpesvirus 6B
Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
Betaherpesvirinae; Roseoloviruses.
REFERENCE
AUTHORS 1 (bases 19696 to 22886)
Pellett,P.E., Sanchez-Martinez,D., Dominguez,G., Black,J.B.,
Anton,E., Greenmoyer,C. and Damhaugh,T.R.
TITLE A strongly immunoreactive vitron protein of human herpesvirus 6
variant B strain Z29: identification and characterization of the
gene and mapping of a variant-specific monoclonal antibody reactive
epitope
JOURNAL Virology 195 (2), 521-531 (1993)
MEDLINE 93331710
PUBMED 7687803
REFERENCE 2 (bases 64106 to 84963)
AUTHORS Stamey,F.R., Dominguez,G., Black,J.B., Damhaugh,T.R. and
Pellett,P.E.
TITLE Intragenomic linear amplification of human herpesvirus 6B orfLyt
suggests acquisition of orfLyt by transposition
JOURNAL J. Virol. 69 (1), 589-596 (1995)
MEDLINE 95074921
PUBMED 7983761
REFERENCE 3 (bases 64106 to 84963; 106012 to 126166)
AUTHORS Lindquester,G.J., Inoue,N., Allen,R.D., Castell,J.W., Stamey,F.R.,
Damhaugh,T.R., O'Brian,J.J., Danovich,R.M., Frenkel,N. and
Pellett,P.E.
TITLE Restriction endonuclease mapping and molecular cloning of the human
herpesvirus 6 variant B strain Z29 genome
JOURNAL Arch. Virol. 141 (2), 367-379 (1996)
MEDLINE 96195263
PUBMED 8634027
REFERENCE 4 (bases 64106 to 84963)
AUTHORS Lindquester,G.J., O'Brian,J.J., Anton,E.D., Greenmoyer,C.A.,
Pellett,P.E. and Damhaugh,T.R.
TITLE Genetic content of a 20.9 kb segment of human herpesvirus 6B strain
Z29 spanning the homologs of human herpesvirus 6A genes U40-57 and
containing the origin of DNA replication
JOURNAL Arch. Virol. 142 (1), 103-123 (1997)

MEDLINE 97300856
PUBMED 9155876
REFERENCE 5 (bases 106012 to 126166)
AUTHORS Lindquester,G.J., Greenmoyer,C.A., Anton,E.D., O'Brian,J.J.,
Pellett,P.E. and Damhaugh,T.R.
TITLE Comparison of a 20 kb region of human herpesvirus 6B with other
human beta herpesviruses reveals conserved replication genes and
adjacent divergent open reading frames
JOURNAL Arch. Virol. 142 (1), 193-204 (1997)
MEDLINE 97300864
PUBMED 9155884
REFERENCE 6 (bases 1 to 162114)
AUTHORS Dominguez,G., Damhaugh,T.R., Stamey,F.R., Dewhurst,S., Inoue,N. and
Pellett,P.E.
TITLE Human herpesvirus 6B genome sequence: coding content and comparison
with human herpesvirus 6A
JOURNAL J. Virol. 73 (10), 8040-8052 (1999)
MEDLINE 99412318
PUBMED 10482553
REFERENCE 7 (bases 19696 to 22886)
AUTHORS Pellett,P.E.
TITLE Direct Submission
JOURNAL Submitted (26-JUL-1993) Herpesvirus Section, Centers for Disease
Control and Prevention, 1600 Clifton Road, Mailstop G18, Atlanta,
GA 30333, USA
REFERENCE 8 (bases 64106 to 84963; 106012 to 126166)
AUTHORS Pellett,P.E.
TITLE Direct Submission
JOURNAL Submitted (02-OCT-1993) Herpesvirus Section, Centers for Disease
Control and Prevention, 1600 Clifton Road, Mailstop G18, Atlanta,
GA 30333, USA
REFERENCE 9 (bases 1 to 162114)
AUTHORS Pellett,P.E., Dominguez,G., Damhaugh,T.R., Stamey,F.R., Dewhurst,S.
and Inoue,N.
TITLE Direct Submission
JOURNAL Submitted (10-JUN-1999) Herpesvirus Section, Centers for Disease
Control and Prevention, 1600 Clifton Road, Mailstop G18, Atlanta,
GA 30333, USA
REMARK Sequence updated by submitter
COMMENT On or before Aug 16 1999 this sequence version replaced gi:405170,
gi:405156, gi:303397.
FEATURES
source Location/Qualifiers
1..162114
/organism="Human herpesvirus 6B"
/mol_type="Genomic DNA"
/strain="Z29"
/db_xref="taxon:32604"
/note="variant B"
1..8793
/rpt_type=direct
59..333
/note="heterogeneous telomeric unit"
/rpt_type=repeat
/rpt_unit="tarsyc"
583..3988
/gene="DR1"
/note="US22 gene family"
join(583..841,955..2975)
/gene="DR1"
/codon_start=1
/product="DR1"
/protein_id="AA049614.1"
/db_xref="GI:5733511"
/translation="MPLTARAGHTLRPLPSHWVLLIGRSLRHVSRYLRGLRL
PLPWPDECHLHPKPYKFLRYPCLTROPHLQGWPDADSLMFDPRYPHPSADSKL
PLGLITSAFMSRVSPTHCSEFHAADPSLMTGSSPMVLADAPGSLFCHDVPFG
RLVLSHSVSLFKTGRROCAIYRABLMRYRPLPSMYTCSDPTATLPLKLARTAR
GLAFLYLMRLHLCGRSLISHPULKEMERTLYVLDPRGMPCCHLISGSSORSSD
AGDTWMAITEKAACGKBAERGGKQCATDRLASHLTRGRDCCRSLOGERSAAEFP
ACRCPILDELTCGEGMPLFATLNSPVACLITLIGGVLAIDLNDVTSLWRIADLELL
LRSLSLISGLRLPLRPSSGSGAARPGYKEKSGRGATYASATATSPRRTPRR
GVTKRGVLTGDPVPSAHPSESECTDGHGQBSGCHDQSGDGRGHRDGGARRHAND
ETEPQGRHEDGQTDGREDDAQESEVARRDEKGTGGGSGRSCGAAOTYGGRGE


```
HGAMSSIPLSVPRPDPRVWVPPHLLFPSPPLPSITPVEDEPSARPRCPGPAEBSKC
SPCPSPDPAPOSAPVRLSALSPSPSPARVRFSLSSSSSSSSSSSSSSSSSSPSPL
SPSPVSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSP
ASPSACICRSRSPSAQTA"
complement (2723..3325)
/gene="DR3"
/evidence=not_experimental
complement (2723..3325)
/gene="DR3"
/evidence=not_experimental
/codon_start=1
/product="DR3"
/protein_id="AAD49618.1"
/db_xref="GI:5733515"
/translation="MSRVFSCVLRACVACGCGCCWCMGYICGDCCRMRRRCARNGRV
GSRVLDGAMRVSGDASVSRTCTERA.PSARSPVSPFVIVSSSSSSSSSSAC
SRRVSPSPSPSHASHAVCAEGGDL.PMGADGDADDEGTGTLERKGADGAGGNAT
GCEPDTHGFARSGLDMGEMNGDLGDEGTGEGDNGEGE"
3022..3988
/gene="B1"
/evidence=not_experimental
3022..3501
/gene="B1"
/codon_start=1
/product="B1"
/protein_id="AAD49617.1"
/db_xref="GI:5733514"
/translation="WQNMKTKTKTKRGRKEGNTPTERRMPAPASRTSAITSGLR
SGSTPLRPGPEVRHA.PTWRTASATTADSHRSPPTPSRRRHITGARTIETS
AAINGVYARAAYVTKRSETIDRLLSVLPGHPRASLRSHLRASALRPDPDR"
3536..3988
/gene="B2"
/evidence=not_experimental
3536..3775
/gene="B2"
/codon_start=1
/product="B2"
/protein_id="AAD49616.1"
/db_xref="GI:5733513"
/translation="MOPMTKKKTKTASKSPARPSLSPPLSRRRVGGSKSERLHGN
APPTPPRSPLPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSP
3912..3917
/note="DR1; B1; B2"
/evidence=not_experimental
3938..3942
/note="DR1; B1; B2"
/evidence=not_experimental
3968..3974
/note="DR1; B1; B2"
/evidence=not_experimental
3983..3988
/note="DR1; B1; B2"
/evidence=not_experimental
5027..7435
/gene="DR6"
/note="US22 gene family; transactivator"
join(5027..5530,6329..7203)
/gene="DR6"
/codon_start=1
/product="DR6"
/protein_id="AAD49615.1"
/db_xref="GI:5733512"
/translation="MTTRHTQTRDGRIRAIRRQGARLHARRARREWILLARGRSKL
YGTSTRHGERIHL.PWRYWCLELHPDPRDARSATVWGHMGWPTTHVRPSVDCA
LDSLYVCCGYGKLOPVGVSSYLT.HSPLDTRLVLLVGRDAVYVHNRARLQRLA
SNTEFARGLQDPVAVEDLEL.PDRMCGNVRHLFPVIAAADHDLTLVGLGICQ
THAGVCELLETETVAVGVRMTLTVARAQYRLMPARQLRLHLIYAGHLGPMI
VCAVLSRERETQSPSPICSGVILGNVPTPDRVETAWVITVAGPLISWPNKG
ICRLANSFALNRMPRARGHWTYSAPGRHLPGDAMPICENVRPVGTLPRQRAYLD
"
7349..7528
/gene="B3"
7349..7528
/gene="B3"
CDS
```

```
polyA_signal
repeat_region
/codon_start=1
/product="B3"
/protein_id="AAD49619.1"
/db_xref="GI:5733516"
/translation="MSIRLRLVCGCYVRYTVLS.PAFNPMRNKIGECTOSHWCRLAV
LIRPHNKKKXHTSGP"
7430..7435
/gene="DR6"
/evidence=not_experimental
8250..8711
/rpt_type=tandem
Query Match 84.0%; Score 16.8; DB 14; Length 162114;
Best Local Similarity 90.0%; Pred. No. 3.1e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 1 AGTCGAACACGTACATGG 20
Db 74437 AGTCGAACACGTACATGG 74456
RESULT 31
ACI04982
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens
Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 164107)
Birtten,B., Linton,L., Nusbaum,C. and Lander E.
Homo sapiens chromosome 17, clone RP11-338L22
2 (bases 1 to 164107)
Birtten,B., Linton,L., Nusbaum,C., Lander,E., All,A., Allen,N.,
Anderson,S., Barna,N., Bastien,V., Boguslavsky,L., Boukhalter,B.,
Brown,A., Camarata,J., Campoliano,A., Chang,J., Chazaro,B.,
Chapel,Y., Colangelo,M., Collins,S., Collamore,A., Cook,A.,
Cooke,P., Dearellano,K., Dewar,K., Diaz,V.S., Dodge,S., Faro,S.,
Ferreira,P., Fitzhugh,W., Gage,D., Galagan,J., Gardyna,S.,
Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
Hagos,B., Heaford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R.,
Jones,C., Kamat,A., Karatas,A., Kells,C., Labocque,K.,
Lamazares,R., Landers,T., Lehoczy,J., Levine,R., Liu,G.,
Maclean,C., MacDonald,P., Major,J., Marquis,N., Matthews,C.,
McCarthy,M., McEwan,P., McKernan,K., McPheters,R., Meldrum,J.,
Menius,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C.,
Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D.,
Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,
Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,
Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,N., Stojanovic,N.,
Seaman,S., Severy,P., Spencer,B., Stange-Thoman,N., Theodorou,J.,
Strauss,N., Subramaniam,A., Talamas,J., Tesfaye,S., Theodorou,J.,
Topham,K., Travers,M., Travis,N., Trigglio,J., Vassiliou,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
Zainoun,D., Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
TITLE
JOURNML
REFERENCE
AUTHORS
```


TITLE
JOURNAL
REFERENCE
AUTHORS

Liou, G., Maclean, C., Macdonald, P., Major, J., Margulis, N.,
Matthews, C., McCarthy, M., McEwan, P., McKenna, K., Meldrum, J.,
Meneus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C.,
Nicoli, R., Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P.,
O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N.,
Pollara, V., Raymond, C., Retra, R., Rieback, M., Riley, R., Rise, C.,
Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S.,
Schuback, R., Seaman, S., Severy, P., Spencer, B., Stange-Thamann, N.,
Stojanovic, N., Straus, N., Subramanian, A., Talamas, J., Tesfaye, S.,
Theodore, J., Topham, K., Travers, M., Travis, N., Triggillo, J.,
Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J.,
Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission
Submitted (20-UN-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
4 (bases 1 to 164107)

Birren, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N.,
Anderson, S., Barna, N., Bastien, V., Bloom, T., Boguslavsky, L.,
Boukhalter, B., Brown, A., Camarata, J., Campopiano, A., Chang, I.,
Chazaro, B., Choepel, Y., Colangelo, M., Collins, S., Collymore, A.,
Cook, A., Cooke, P., DeArillano, K., Dewar, K., Diaz, J.S., Dodge, S.,
Faro, S., Ferreira, P., Fitzgerald, M., Fitzhugh, W., Gage, D.,
Galagan, J., Gardina, S., Ginde, S., Gord, S., Goyette, M., Graham, L.,
Grand-Pierre, N., Hages, B., Horton, L., Hulme, W., Iliev, I.,
Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., Labrecque, K.,
Lamarez, R., Landers, T., Lehocaky, J., Levine, R., Lindblad-Toh, K.,
Liou, G., Maclean, C., Macdonald, P., Major, J., Margulis, N.,
Matthews, C., McCarthy, M., McEwan, P., McKenna, K., Meldrum, J.,
Meneus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C.,
Nicoli, R., Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P.,
O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N.,
Pollara, V., Raymond, C., Retra, R., Rieback, M., Riley, R., Rise, C.,
Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S.,
Schuback, R., Seaman, S., Severy, P., Spencer, B., Stange-Thamann, N.,
Stojanovic, N., Straus, N., Subramanian, A., Talamas, J., Tesfaye, S.,
Theodore, J., Topham, K., Travers, M., Travis, N., Triggillo, J.,
Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J.,
Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

TITLE
JOURNAL
COMMENT

Direct Submission
Submitted (24-UN-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Jun 20, 2002 this sequence version replaced g1:21431101.
All repeats were identified using RepeatMasker:
Smith, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: <http://www-seq.wi.mit.edu>
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L2281
Center clone name: 338_L_22

FEATURES
source
----- Location/Qualifiers
1.164107
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="17"
/map="17"
/clone="RP11-338L22"
/clone_lib="RPC1-11 Human Male BAC"
repeat_region
/rpt_family="AluA"
complement(245..426)
repeat_region
/rpt_family="Alu"
complement(427..738)
repeat_region
/rpt_family="AluX"
complement(739..1453)
repeat_region
/rpt_family="Alu"
1454..1703
/rpt_family="TTCOC)n"

unsure
/note="single clone coverage"
1657..1660
unsure
/note="probably T, possibly TT"
1665
complement(1704..1894)
repeat_region
/rpt_family="L1"
1907..2021
repeat_region
/rpt_family="L1M"
2096..2233
repeat_region
/rpt_family="L1M3A"
complement(2286..2511)
repeat_region
/rpt_family="L1M4"
3574..3747
repeat_region
/rpt_family="AluY"
3782..3824
repeat_region
/rpt_family="AT_rich"
3927..4011
repeat_region
/rpt_family="L2"
4014..4265
repeat_region
/rpt_family="AluX"
4747..5057
repeat_region
/rpt_family="AluX"
5553..7200
repeat_region
/rpt_family="MER52A"
7354..7527
repeat_region
/rpt_family="MER5B"
7674..7791
repeat_region
/rpt_family="MIR"
complement(8537..8741)
repeat_region
/rpt_family="MIR"
8923..9081
repeat_region
/rpt_family="L1ME3A"
complement(9482..9805)
repeat_region
/rpt_family="AluO"
complement(9811..9943)
repeat_region
/rpt_family="MER5B"
9944..10243
repeat_region
/rpt_family="AluX"
complement(10244..10287)
repeat_region
/rpt_family="MER5B"
11025..11346
repeat_region
/rpt_family="AluX"
12802..12858
repeat_region
/rpt_family="L2"
complement(13372..13631)
repeat_region
/rpt_family="AluO"
complement(13813..14105)
repeat_region
/rpt_family="AluX"
15481..15908
repeat_region
/rpt_family="MER6A"
16970..17272
repeat_region
/rpt_family="AluY"
17489..17628
repeat_region
/rpt_family="AluX"
17629..17853
repeat_region
/rpt_family="(TA)n"
17654..17822
repeat_region
/rpt_family="AluX"
18478..18780
repeat_region
/rpt_family="AluY"
18877..18898
repeat_region
/rpt_family="AT_rich"
19087..19398
repeat_region
/rpt_family="AluG"
19411..19655
repeat_region
/rpt_family="AluO"
19772..20060
repeat_region
/rpt_family="AluX"
complement(20327..20621)
repeat_region
/rpt_family="AluY"
20742..20833
repeat_region
/rpt_family="L1M4"
complement(20859..21156)

repeat_region /rpt_family="AluSc"
21414..21583
/rpt_family="MIR"
repeat_region 21592..21739

Query Match 84.0%; Score 16.8; DB 9; Length 164107;
Best Local Similarity 90.0%; Pred. No.3.1e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AGTCGCAACAGTACATGG 20
|||||
Db 119533 AGTCACAAACAGTACAGGG 119552

RESULT 32
AC083894/c 167273 bp DNA linear ROD 14-NOV-2002
LOCUS AC083894
DEFINITION Mus musculus chromosome 6 clone rp23-259j8 strain C57BL/6J,
complete sequence.
ACCESSION AC083894.21 GI:24961434
VERSION
KEYWORDS HTG.
SOURCE Mus musculus (house mouse)
ORGANISM

REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
TITLE 1 (bases 1 to 167273)
AUTHORS Loh, P., Qi, S., Ford, B. and Roe, B.A.
JOURNAL Mus musculus BAC clone rp23-259j8
UNPUBLISHED
2 (bases 1 to 167273)
AUTHORS Loh, P., Qi, S., Ford, B. and Roe, B.A.
JOURNAL Direct Submission
TITLE Submitted (05-OCT-2000) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
REFERENCE
AUTHORS 3 (bases 1 to 167273)
TITLE Loh, P., Qi, S., Ford, B. and Roe, B.A.
JOURNAL Direct Submission
TITLE Submitted (21-JUL-2002) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA

REFERENCE
AUTHORS 4 (bases 1 to 167273)
TITLE Loh, P., Qi, S., Ford, B. and Roe, B.A.
JOURNAL Direct Submission
TITLE Submitted (26-JUL-2002) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
REFERENCE
AUTHORS 5 (bases 1 to 167273)
TITLE Loh, P., Qi, S., Ford, B. and Roe, B.A.
JOURNAL Direct Submission
TITLE Submitted (03-AUG-2002) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
REFERENCE
AUTHORS 6 (bases 1 to 167273)
TITLE Loh, P., Qi, S., Ford, B. and Roe, B.A.
JOURNAL Direct Submission
TITLE Submitted (14-NOV-2002) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA

COMMENT On Nov 14, 2002 this sequence version replaced gi:22094402.

----- Genome Center
Center: Department Of Chemistry And Biochemistry
The University Of Oklahoma
Center code:UOKXOR

FEATURES
source Location/Qualifiers
1..167273
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/chromosome="6"

ORIGIN /clone="rp23-259j8"
/clone_id="RP01 - 23 Female (C57BL/6J) Mouse BAC Library"

Query Match 84.0%; Score 16.8; DB 10; Length 167273;
Best Local Similarity 90.0%; Pred. No.3.1e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AGTCGCAACAGTACATGG 20
|||||
Db 5972 AGTTGAAACAGTACATGG 5953

RESULT 33
AC129943 178546 bp DNA linear HTG 06-AUG-2004
LOCUS AC129943
DEFINITION Mus musculus chromosome 5 clone RP23-410W21 map 5, *** SEQUENCING
IN PROGRESS ***.
ACCESSION AC129943.7 GI:51011211
VERSION
KEYWORDS HTG; HTGS PHASE2; HTGS FULLTOP; HTGS_ACTIVEFIN.
SOURCE Mus musculus (house mouse)
ORGANISM

REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
TITLE 1 (bases 1 to 178546)
AUTHORS Birren, B., Nusbaum, C. and Lander, E.
JOURNAL Mus musculus chromosome 5, clone RP23-410W21
UNPUBLISHED
2 (bases 1 to 178546)
AUTHORS Birren, B., Nusbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S.,
Barna, N., Bastien, V., Bloom, T., Boguslavsky, J., Boukhgalter, B.,
Camarata, J., Chang, J., Chazaro, B., Choepel, Y., Collymore, A.,
Cook, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J.S., Dodge, S.,
Faro, S., Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J.,
Gardyna, S., Gord, S., Graham, L., Grand-Pierre, N., Hagos, B.,
Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A.,
Karatas, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K.,
Liu, G., Maclean, C., Macdonald, P., Major, J., Matthews, C.,
McCarthy, M., Meldrum, J., Menus, L., Mihova, T., Mlenga, V.,
Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C.H.,
O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K.,
Phunkhang, P., Pierre, N., Raymond, C., Retz, R., Rise, C., Rogov, P.,
Roman, J., Roy, A., Schauer, S., Schnupack, R., Seaman, S., Severy, P.,
Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Talamas, J.,
Tesfaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H.,
Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J.,
Zembek, L., Zimmer, A. and Zody, M.

TITLE Direct Submission
JOURNAL Submitted (06-AUG-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 178546)
AUTHORS Birren, B., Nusbaum, C., Lander, E., Aboueleil, A., Allen, N.,
Anderson, M., Anderson, S., Arachchi, H.M., Barna, N., Bastien, V.,
Bloom, T., Boguslavsky, J., Boukhgalter, B., Camarata, J., Chang, J.,
Choepel, Y., Collymore, A., Cook, A., Cooke, P., Corum, B.,
Dearellano, K., Diaz, J.S., Dodge, S., Dooley, K., Dorris, L.,
Erickson, J., Faro, S., Ferreira, P., Fitzgerald, M., Gage, D.,
Galagan, J., Gardyna, S., Graham, L., Grand-Pierre, N., Hagos, B.,
Hagopian, D., Hagos, B., Hall, J., Horton, L., Hulme, W., Iliev, I.,
Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., Landers, T.,
Levine, R., Lindblad-Toh, K., Liu, G., Liu, X., Lui, A., Mabbitt, R.,
McLean, C., Macdonald, P., Major, J., Manning, J., Matthews, C.,
McCarthy, M., Meldrum, J., Menus, L., Mihova, T., Mlenga, V.,
Murphy, T., Naylor, J., Nguyen, C., Nguyen, T., Nicol, R., Norbu, C.,
O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K.,
Phunkhang, P., Pierre, N., Rachupka, A., Ramasamy, J., Raymond, C.,
Retz, R., Rise, C., Rogov, P., Roman, J., Schauer, S., Schnupack, R.,
Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N.,
Stojanovic, N., Stubbs, M., Talamas, J., Tesfaye, S., Theodore, J.,
Topham, K., Travers, M., Vassiliev, H., Venkataraman, V.S., Viel, R.,
Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L.,
Zimmer, A. and Zody, M.

TITLE Direct Submission
JOURNAL Submitted (06-AUG-2004) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT On Aug 6, 2004 this sequence version replaced gi:50511617.
All repeats were identified using RepeatMasker:
Smith, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/MIT Center for Genome Research
Center code: WITBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@road.mit.edu
----- Project Information
Center project name: 126468
Center clone name: 410_M21

* NOTE: This is a 'working draft' sequence. It currently
* consists of 1 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
* 1 178546: contig of 178546 bp in length.
Location/Qualifiers
1. 178546
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="5"
/map="5"
/clone="RP23-410M21"
/clone_1ib="RP23-410M21"
/clone_1ib="RP23-410M21"

ORIGIN
Query Match 84.0%; Score 16.8; DB 2; Length 178546;
Best Local Similarity 90.0%; Pred. No. 3.1e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AGTCGCAACAGTACATGG 20
|||||
Db 157450 AGTCGCAACAGTACATGG 157469

RESULT 34
AC140318 180212 bp DNA linear HTG 14-JUL-2004
LOCUS Mus musculus chromosome 9 clone RP23-400F19, WORKING DRAFT
DEFINITION
SEQUENCE
AC140318
AC140318.3 GI:50284668
HTG; HTGS PHASE1; HTGS DRAFT; HTGS_ACTIVEPERIN.
KEYWORDS
SOURCE Mus musculus (house mouse)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 180212)
Wilson, R.K.
The sequence of Mus musculus clone
Unpublished
2 (bases 1 to 180212)
McPherson, J.D. and Waterston, R.H.
Direct Submission
Submitted (23-FEB-2003) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA
3 (bases 1 to 180212)
Wilson, R.K.
Direct Submission
Submitted (14-JUL-2004) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA
JOURNAL
TITLE
AUTHORS
COMMENT On Jul 14, 2004 this sequence version replaced gi:29824550.

----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu
Contact: submissions@wustl.wustl.edu
----- Project Information -----
Center project name: M BA0400F19
----- Summary Statistics -----
Sequencing vector: M13; 0%
Sequencing vector: plasmid; 100%
Chemistry: Dye-Primer ET; 0% of reads
Chemistry: Dye-Terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 180137 bases at least Q40
Consensus quality: 180190 bases at least Q30
Consensus quality: 180203 bases at least Q20

* NOTE: This is a 'working draft' sequence. It currently
* consists of 1 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 180212: contig of 180212 bp in length.
Location/Qualifiers
1. 180212
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="5"
/clone="RP23-400F19"
/clone="RP23-400F19"
/note="assembly_name:Contig24"

ORIGIN
Query Match 84.0%; Score 16.8; DB 2; Length 180212;
Best Local Similarity 90.0%; Pred. No. 3.1e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AGTCGCAACAGTACATGG 20
|||||
Db 53960 AGTCGCAACAGTACATGG 53941

RESULT 35
AC147514 187725 bp DNA linear ROD 06-AUG-2004
LOCUS Mus musculus chromosome 1, clone RP23-442A6, complete sequence.
DEFINITION
SEQUENCE
AC147514
AC147514.11 GI:51011231
HTG.
KEYWORDS
SOURCE Mus musculus (house mouse)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 187725)
Birren, B., Nusbaum, C. and Lander, E.
Mus musculus chromosome 1, clone RP23-442A6
Unpublished
2 (bases 1 to 187725)
Birren, B., Nusbaum, C., Lander, E., Aboueleil, A., Allen, N.,
Anderson, M., Arachchi, H.M., Barna, N., Bastien, V., Bloom, T.,
Boguslavsky, L., Boudgallier, B., Camarero, V., Chang, V., Choepel, Y.,
Collymore, A., Cook, A., Cooke, P., Corum, B., DeArrelano, K.,
Diaz, V.S., Dodge, S., Dooley, K., Dorris, L., Erickson, J., Faro, S.,
Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J., Gardys, S.,
Graham, L., Grand-Pierre, N., Hafez, N., Hagopian, D., Hago, B.,
Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,
Kamati, A., Karates, A., Kells, C., Landers, T., Levine, R.,


```

repeat_region      20507..21000
                    /rpt_family="LxS"
repeat_region      21070..21104
                    /rpt_family="AT_rich"
repeat_region      21430..21563
                    /rpt_family="AT_rich"
repeat_region      22581..22651
                    /rpt_family="B1F"
repeat_region      22660..22687
                    /rpt_family="(A)n"
repeat_region      complement(22688..23503)
                    /rpt_family="RMR1A"
repeat_region      complement(23478..23989)
                    /rpt_family="RMR1A"
repeat_region      23992..24050
                    /rpt_family="(TTC)n"
repeat_region      complement(24051..24122)
                    /rpt_family="L1_MM"
repeat_region      24123..24146
                    /rpt_family="(CACC)n"
repeat_region      complement(24147..24528)
                    /rpt_family="L1_MM"
repeat_region      24552..25808
                    /rpt_family="Lx"
repeat_region      25809..25912
                    /rpt_family="ORR1B-int"
repeat_region      25913..26317
                    /rpt_family="ORR1C"
repeat_region      26318..26399
                    /rpt_family="ORR1D"
repeat_region      27209..27235

```

Query Match 84.0%; Score 16.8; DB 10; Length 187725;

Best Local Similarity 90.0%; Pred. No. 3.1e+02; Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AGTCGCAACAGTACATGG 20
 Db 69092 AGTCGCAACAGTACATGG 69111

RESULT 36 AC079443 207091 bp DNA linear ROD 19-NOV-2003
 LOCUS AC079443 Mus musculus clone rp23-381f7 map 6 strain C57BL/6J, complete

DEFINITION Mus musculus clone rp23-381f7 map 6 strain C57BL/6J, complete

ACCESSION AC079443
 VERSION AC079443
 KEYWORDS HTG.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus

REFERENCE 1 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 207091)
 Q1.S., Loh,P., Ford,B. and Roe,B.A.

REFERENCE 2 Mus musculus BAC Clone rp23-381f7
 2 (bases 1 to 207091)
 Q1.S., Loh,P., Ford,B. and Roe,B.A.
 Direct Submission
 Submitted (01-SEP-2000) Department Of Chemistry And Biochemistry,
 The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
 OK 73019, USA

REFERENCE 3 (bases 1 to 207091)
 Q1.S., Loh,P., Ford,B. and Roe,B.A.
 Direct Submission
 Submitted (08-NOV-2003) Department Of Chemistry And Biochemistry,
 The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
 OK 73019, USA

REFERENCE 4 (bases 1 to 207091)
 Q1.S., Loh,P., Ford,B. and Roe,B.A.
 Direct Submission
 Submitted (18-NOV-2003) Department Of Chemistry And Biochemistry,

REFERENCE 5 (bases 1 to 207091)
 Q1.S., Loh,P., Ford,B. and Roe,B.A.
 Direct Submission
 Submitted (19-NOV-2003) Department Of Chemistry And Biochemistry,
 The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
 OK 73019, USA

COMMENT On Nov 18, 2003 this sequence version replaced gi:3822376.

----- Genome Center
 Center: Department Of Chemistry And Biochemistry
 The University Of Oklahoma
 Center code:UOKNOR

FEATURES
 source Location/Qualifiers
 1..207091
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /map="6"
 /clone="rp23-381f7"
 /clone_id="RPCT - 23 Female (C57BL/6J) Mouse BAC Library"

ORIGIN
 Query Match 84.0%; Score 16.8; DB 10; Length 207091;
 Best Local Similarity 90.0%; Pred. No. 3.1e+02;
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AGTCGCAACAGTACATGG 20
 Db 19583 AGTCGCAACAGTACATGG 19602

RESULT 37 AC132022 229916 bp DNA linear HTG 15-NOV-2002
 LOCUS AC132022/c Rattus norvegicus clone CH230-490L1, *** SEQUENCING IN PROGRESS

DEFINITION *** 5 unordered pieces.

ACCESSION AC132022
 VERSION AC132022.3 GI:25007381
 KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED.
 SOURCE Rattus norvegicus (Norway rat)
 ORGANISM Rattus norvegicus

REFERENCE 1 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 1 (bases 1 to 229916)
 Muzny,D.Marie, Metzker,M.Lee, Abramson,S., Adams,C., Alder,J., Allen,C., Allen,H., Alsbrooks,S., Amiri,A., Angiano,D., Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H., Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Bernhamed,F., Biswal,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M., Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E., Cardenas,V., Carter,K., Cavazos,I., Caesar,H., Center,A., Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J., Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L., Davila,M.B., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D., Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K., Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K., Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G., Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P., Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M., Gebregorgis,E., Geer,K., Gill,R., Grady,M., Guerra,N., Guetara,W., Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K., Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J., Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hognes,M., Hollins,B., Howells,S., Huylk,S., Hume,J., Idlebird,D., Jackson,A., Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A., Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C., Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J., Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J., Lorenshewa,L., Loulseged,H., Lozado,R.J., Lu,X., Ma,J.,

Maheshwari, M., Mahindartne, M., Mahmoud, M., Mallory, K., Mangum, A., Mangum, B., Mapa, P., Martin, K., Martin, R., Martinez, E., McWhinney, S., McLeod, M. P., McNeill, T. Z., Meenen, E., Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., O'Brien, O., Okonofu, G., Olariu, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C., Plopper, F., Polidexter, A., Popovic, D., Primus, E., Pu, L., L., Puzo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M. A., Relph, R., Kelly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S. J., Sanders, W., Savary, G., Scherer, S., Scott, G., Shatsman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C. D., Smajda, D., Sneed, A., Sodergren, E., Song, X. Z., Sorelle, R., Sosa, D., Steinhilber, M., Strong, R., Sutton, A., Svatek, A., Taber, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Uman, K., Valas, R., Vera, V., Villalana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wleczek, R., Wooden, H., Wortley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, Y., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Niederhausen, A., Weiss, R., Smith, D. R., Holt, R. A., Smith, H. O., Weinstein, G., and Gibbs, R. A.

Unpublished
2 (bases 1 to 229916)
Rat Genome Sequencing Consortium.
Direct Submission
Submitted (29-AUG-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 229916)
Rat Genome Sequencing Consortium.
Direct Submission
Submitted (15-NOV-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
On Nov 15, 2002 this sequence version replaced gi:2398476.
The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu

----- Project Information
Center project name: KBRB
Center clone name: CH230-490L11

----- Summary Statistics
Assembly program: Phrap, version 0.990329
Consensus quality: 192103 bases at least Q40
Consensus quality: 195577 bases at least Q30
Consensus quality: 197895 bases at least Q20
Estimated insert size: 198652; sum-of-contigs estimation
Quality coverage: 5x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_drift_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 5 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as

* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 23162: contig of 23162 bp in length
* 23163 23262: gap of unknown length
* 23263 190541: contig of 167279 bp in length
* 190542 190641: gap of unknown length
* 190642 203791: contig of 13150 bp in length
* 203792 203891: gap of unknown length
* 203892 222968: contig of 19077 bp in length
* 222969 223068: gap of unknown length
* 223069 223162: gap of 6848 bp in length.
Location/Qualifiers
1. .229916
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-490L11"
1. .3229
/note="wgs_end_extension
clone_end:Sp6"
6915. .8054
/note="clone boundary
clone_end:Sp6"
end_sequence="B2128304"
23263. .24747
/note="wgs_contig"
complement(30107. .31256)
/note="clone boundary
clone_end:T7
site:
end_sequence="B2128302"
44736. .46308
/note="wgs_end_extension
clone_end:T7"
47013. .49058
/note="wgs_end_extension
clone_end:T7"
190642. .199974
/note="wgs_end_extension
clone_end:T7"
200388. .201787
/note="wgs_end_extension
clone_end:T7"
201835. .203791
/note="wgs_end_extension
clone_end:T7"
203892. .211452
/note="wgs_end_extension
clone_end:T7"
213252. .215045
/note="wgs_end_extension
clone_end:T7"
217112. .218575
/note="wgs_end_extension
clone_end:T7"
218820. .220272
/note="wgs_end_extension
clone_end:T7"
220774. .222968
/note="wgs_end_extension
clone_end:T7"

ORIGIN
Query Match 84.0%; Score 16.8; DB 2; Length 229916;
Best Local Similarity 90.0%; Pred. No. 36-02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 AGTCGCAACAGTACATG 20
Db 142955 AGTCACAAACAGTACATG 142936

RESULT 38
AC112824
LOCUS
DEFINITION
AC112824 clone CH230-365H1, *** SEQUENCING IN PROGRESS

AC112824 GI:24635565
HTG: HTGS PHASE2; HTGS DRAFT; HTGS_ENRICHED.
Rattus norvegicus (Norway rat)
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 240243)
Muzny, D., Maric, M., Metzger, M., Lee, A., Abramson, S., Adams, C., Alder, J.,
Allen, C., Allen, H., Alsbrooks, S., Amth, A., Anguiano, D.,
Anyalebech, V., Ayagi, A., Ayodeji, M., Baca, E., Baden, H.,
Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F.,
Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M.,
Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E.,
Cardenas, V., Carter, K., Cavazos, I., Caesar, H., Center, A.,
Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J.,
Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L.,
Davila, M., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D.,
Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K.,
Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K.,
Egan, A., Escoto, M., Eugene, C., Evans, C.A., Falls, T., Fan, G.,
Fernandez, S., Finley, M., Flaggs, N., Forbes, L., Foster, M., Foster, P.,
Fraser, C.M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M.,
Gibbs, G., Gish, A., Gill, R., Grady, M., Guerra, M., Guevara, W.,
Gunnarsson, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K.,
Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J.,
Hernandez, R., Hines, S., Hladun, S.L., Hodgson, A., Hognes, M.,
Hollins, B., Howells, S., Huik, S., Hume, J., Idlebird, D., Jackson, A.,
Jacks, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolyet, A.,
Karpathy, S., Kelly, S., Kelly, S., Khan, Z., King, J., Kovar, C.,
Kowis, C., Kraft, C.L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J.,
Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J.,
Lorensueta, L., Louie, H., Lozano, R.J., Lu, X., Ma, U.,
Maheshwari, M., Mahindaratne, M., Mahmoud, M., Vailley, K., Mangum, A.,
Mangum, B., Mapa, P., Martin, K., Martin, R., Martinez, E.,
Mawhinney, S., McLeod, M.P., McNeill, T.Z., Meenen, E.,
Mlosoavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S.,
Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L.,
Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Parks, K.,
Pasacki, O., Okwuonu, G., Olarunpusagun, A., Pal, S., Parks, K.,
Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C.,
Plopper, F., Polindexter, A., Popovic, D., Primus, E., Pu, L.,
Pozo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M.A., Reigh, R.,
Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F.,
Rivers, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S.J.,
Sanders, W., Saverly, G., Scherer, S., Scott, G., Shatsman, S., Shear, H.,
Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C.D., Smajs, D.,
Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J.,
Steinle, M., Strong, R., Sutton, A., Syatek, A., Tabor, P., Taylor, C.,
Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K.,
Valas, R., Vera, V., Villaseca, D., Waldron, L., Walker, B., Wang, J.,
Wang, O., Wang, S., Warren, J., Warren, R., Woden, H., Morley, K.,
Williams, G., Willson, R., Wlezyk, R., Wodden, H., Morley, K.,
Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V.,
Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von
Niederhausen, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O.,
Weinstock, G. and Gibbs, R.A.
Direct Submission
Unpublished
2 (bases 1 to 240243)
Morley, K.C.
Direct Submission
Submitted (25-FEB-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
3 (bases 1 to 240243)
Rat Genome Sequencing Consortium.
Direct Submission
Submitted (19-NOV-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Nov 6, 2002 this sequence version replaced gi:23603863.
The sequence in this assembly is a combination of BAC based reads
and whole genome shotgun sequencing reads assembled using Atlas
(http://www.hgsc.bcm.tmc.edu/projects/rac/). Each contig described
in the feature table below represents a scaffold in the Atlas
assembly (a 'contig-scaffold'). Within each contig-scaffold,
individual sequence contigs are ordered and oriented, and separated
by sized gaps filled with Ns to the estimated size. The sequence
may extend beyond the ends of the clone and there may be sequence
contigs within a contig-scaffold that consist entirely of whole
genome shotgun sequence reads. Both end sequences and whole genome
shotgun sequence only contigs will be indicated in the feature
table.
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GRER
Center clone name: CH230-365H1
----- Summary Statistics
Assembly program: Phrap; version 0.990329
Consensus quality: 168357 bases at least Q40
Consensus quality: 170011 bases at least Q30
Consensus quality: 170942 bases at least Q20
Estimated insert size: 175728; sum-of-contigs estimation
Quality coverage: 6x in Q20 bases; sum-of-contigs estimation
* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 1 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
1 240243: contig of 240243 bp in length.
Location/Qualifiers
1. 240243
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-365H1"
1. 1773
/note="wgs_contig"
ORIGIN
Query Match 84.0%; Score 16.8; DB 2; Length 240243;
Best Local Similarity 90.0%; Pred. No. 3e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 AGTCGCAACAGTACATCG 20
Db 193218 AGTCACAAACAGTACATCG 193237
RESULT 39
AC120133
LOCUS
DEFINITION
IN PROGRESS
AC120133 GI:39841417
282567 bp DNA linear HTG 15-DEC-2003
Mus musculus chromosome 6 clone RP23-11L19 map 6, *** SEQUENCING

TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
AUTHORS
TITLE
JOURNAL
COMMENT

Submitted (09-NOV-2001) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
(bases 1 to 321695)
Rat Genome Sequencing Consortium.
Submitted (10-MAY-2003) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
On May 10, 2003 this sequence version replaced gi:24819582.
The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be whole contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
Project Information
Center project name: GHCM
Center clone name: CH230-51C23
----- Summary Statistics
Assembly program: Atlas 3.0;
Consensus quality: 261123 bases at least Q40
Consensus quality: 264929 bases at least Q30
Consensus quality: 267345 bases at least Q20
Estimated insert size: 281114; sum-of-contigs estimation
Quality coverage: 6x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html)
* NOTE: This sequence may represent more than one clone.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 8 contigs. The true order of the pieces

is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1 24448: contig of 24448 bp in length
24449 24548: gap of unknown length
24549 43107: contig of 18559 bp in length
43107 43207: gap of unknown length
43207 43208: contig of 100 bp in length
43208 47877: gap of 4669 bp in length
47877 47977: gap of unknown length
47977 312827: gap of 264851 bp in length
312827 312928: gap of unknown length
312928 314841: contig of 1913 bp in length
314841 314940: gap of unknown length
314940 314941: contig of 1499 bp in length
314941 316539: gap of unknown length
316539 316540: contig of 2094 bp in length
316540 318634: gap of unknown length
318634 318734: gap of unknown length
318734 321695: contig of 2962 bp in length.
Location/Qualifiers
1..321695
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-51C23"
1..1526
/note="wgs_end_extension
clone_end:5p6"
complement(6233..6853)
/note="clone boundary
clone_end:5p6
site:ECORI
end_sequence:BH337792"
15203..24448
/note="wgs_contig"
43208..44716
/note="wgs_contig"
45034..47876
/note="wgs_contig"
47977..49325
/note="wgs_contig"

ORIGIN
Query Match 84.0%; Score 16.8; DB 2; Length 321695;
Best Local Similarity 90.0%; Pred. No. 3e+02; 2; Indels 0; Gaps 0;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 AGTCGCAACAGTACATG 20
Db 273918 AGTCACAAACGTAGATGG 273937

RESULT 41
BX571870 348505 bp DNA linear BCT 26-SEP-2003
LOCUS
DEFINITION Phocorhabdus luminescens subsp. launondii T101 complete genome;
segment 12/17.
ACCESSION BX571870 BX470251
VERSION BX571870.1 GI:36786547
KEYWORDS complete genome.
SOURCE Phocorhabdus luminescens subsp. launondii T101
ORGANISM Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
Enterobacteriaceae; Phocorhabdus.
REFERENCE
AUTHORS
1 Duchaud E., Rusniok C., Frangeul L., Buchrieser C., Taourit S.,
Bocs S., Bouraux-Bude C., Chandler M., Dassa E., Derose R.,
Derzelle S., Freysinet G., Gaudriault S., Givaudan A., Glaeser P.,
Medigne C., Lanois A., Powell K., Sigulier P., Wingate V.,
Zouine M., Boemare N., Danchin A. and Kunst F.
Complete genome sequence of the entomopathogenic bacterium

JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL

Photobabds luminescens
Nat. Biotechnol. 11 (1) (2003) In press
Duchaud, F., Frangeul, L., Rusniok, C. and Kunst, F.
Direct Submission
Submitted (23-APR-2003) L. Frangeul, Institut Pasteur, Genopole, 25
rue du Docteur Roux, 75724 Paris Cedex 15, FRANCE. E-mail:
lfrangeul@pasteur.fr, fkunst@pasteur.fr

FEATURES

source

1. .348505

/organism="Photobabds luminescens subsp. laumondii T701"

/mol_type="genomic DNA"

/strain="T701"

/db_xref="taxon:243265"

complement (477. .3848)

/locus_tag="plu3231"

complement (477. .3848)

/locus_tag="plu3231"

/function="Unknown"

/note="unnamed protein product; Similar to putative
membrane protein of Y. pestis"

/codon_start=1

/transl_table=11

/protein_id="CAE15605.1"

/db_xref="GI:36786548"

/translation="MKVNRHQYGFYIFITLSAGIFLMILYNOQLLEERGJLKS
VWYAVATVPILILGLFENAVWMHTRONKQKRLPEPETERADSSOGALVSDIGDR
MLRGRFWMYKRIILVVEVQVAYIAPGLTTMOGYRTLLMGSLQAEPTA
QELALRLRRYPRINAIWALNENQAOVPMMDKALRMOKOOLRQAPYLMQVC
HSTMGOEGRITOTVGCPEPPECPDPIVATOLITPLBERGMOOLLEKTAEDICRL
SANTKOGGIAHNOQVTPWLAEGDVYSIRGLMFSPLKQCTTRLPQIMPTANQCV
LESSRCFHGRVGMHTQTVCRLIVGLKMGAMLSFTIRDOALAVQAAVDLKY
QENSDAOQMALKDLENEVDRLQRAHEGVNPERFGLSQNKLLTALPDYAVNNR
LARDKAAVYLAHLSALVNLPGSPPLARQAQGHQKAYLMMARPEKVAFLTR
LDNEBERAGIAPGVLTIDAMLSFYENLAHPBMTIADKALIKOVROLLOPG
KHVAEALYQOMOSYKNSDILRQMGTDAAALFTTSHVIGMFTROMEGGIO
KIAKAVASRKALIDWVLSDNROAVASVSPPEELKARLTARVETDPARVMDPMSLN
KAKTRLSVIVQILITMADGROSPLIANTALAYOGETGQCPSSDSIVSAKTLIN
KTPRINIOQIHGPMDETFGPLALMGENQAGRIADSTNLQCFLRVTVDLR
KIQOMNSDDPPMANALAKTVFEGKSTDLTDQYGRILASLGQWRSFGQTLFVE
PLAQMVGVLKPKATNLNSQMTATVDMWDRAPAGYFAGGSEISLIPMGQFIRSD
SGRIEPLTRQLGILHKGAGNVVDEVNGQVQINPDLKALINLSQSLLEFADGS
OGRIEPLRARAVDVEYEDTLTDGQKLRFNEMESQSPRMFGETKRGWMLTMSVN
AGARLGGYQGMWGLIRMLAKARLDESRRLFTAPDGLPLWILRTLEGSPILA
LLKRGFKLPKNTFYVMGENTTISAIVDDDLIEE"

gene

complement (3845. .4924)

/locus_tag="plu3232"

complement (3845. .4924)

/locus_tag="plu3232"

/function="Unknown"

/note="unnamed protein product; Highly similar to unknown
protein of Photobabds luminescens"

/codon_start=1

/transl_table=11

/protein_id="CAE15606.1"

/db_xref="GI:36786549"

/translation="MSPLRRLMFIILLIVGLVYAVLKDITTEVHAIVSSLP
FVILLRGAIVRYEYCAAYAMEEASAEPRFOWQOMRSKALVINCILLTEPQOC
IAPUGDAEIRIAPPEQCRALSSDSRQLAKNDQPEQCCPEYRSLAQITIASD
DHASVESATYQWKISILPILNRDEIDLPDEDEGLIMVIAFCWMSSENNKAYAS
FVSQALFSTKRTAQEKLVLAGMGRTPLSJGELIKDQILFEYNRIDKRNHMLI
TGVIVDTLTELINSHDSAFNLFTENSVHLIDHTGPPGPISIFLASALTEALSLG
SHIIVNQLPEGSASLYMTKEIHQ"

gene

complement (3856. .3861)

/locus_tag="plu3231"

complement (3856. .3861)

/locus_tag="plu3231"

/locus_tag="plu3231"

complement (4944. .5191)

/locus_tag="plu3233"

complement (4944. .5183)

/locus_tag="plu3233"

/function="Unknown"

/note="unnamed protein product; Hypothetical gene"

RBS

gene

CDS

/codon_start=1

/transl_table=11

/protein_id="CAE15607.1"

/db_xref="GI:36786550"

/translation="MIMFMIGFSHITDQSRIVSYALAFYDKKQEVKIGSLIFILN
AKQGMHMYDVCCCPSPIDPRLISSYYELAH"

complement (5186. .5191)

/locus_tag="plu3233"

complement (5196. .5429)

/locus_tag="plu3234"

complement (5196. .5429)

/locus_tag="plu3234"

/function="Unknown"

/note="unnamed protein product; Similar to unknown protein
of Photobabds and some similarities with unknown
protein"

/codon_start=1

/transl_table=11

/protein_id="CAE15608.1"

/db_xref="GI:36786551"

/translation="MGVDMHIFELQVVLGIYLLLYTTPILIDVFTLVRIKKCI
KDKVGRASRLIPAVFMSCSLIAFIWNLN"

complement (5517. .6623)

/locus_tag="plu3235"

complement (5517. .6623)

/locus_tag="plu3235"

/function="Unknown"

/note="unnamed protein product; Highly similar to unknown
protein of Photobabds"

/codon_start=1

/transl_table=11

/protein_id="CAE15609.1"

/db_xref="GI:36786552"

/translation="MNRKTLITLAVICTLMMVMPVYPIITLKEKKTTEMLVMQ
PHCVGRILIDLPAGNVAEOGFIQDAQIETERYLPAPFQRIIRLSEOLTKMOVE
PKMPPLKAYARLDNMGEGVIFPNBSFGIPGPARVLAHLVSNVAFVTKMELS
DPCYKEDDPVYTCGSRSEYQNEIYOTLKKMRLSRISGRDTEIPTETGTCIPCF
ITGSHDKEMMTFYVRKDSDEHTEFVETLNDIOEKHLEIRGHEGFLSMKXITI
RKRRKINGTYIEVLTATPTEKIEEPETKIQYKTLTANETIGDYKNPFVIRNLRN
DRSAPYSSENEIKFWDVSTFRPRPAFRQ"

gene

complement (6620. .8236)

/locus_tag="plu3236"

complement (6620. .8236)

/locus_tag="plu3236"

/function="Unknown"

/note="unnamed protein product; Similar to unknown
protein"

/codon_start=1

/transl_table=11

/protein_id="CAE15610.1"

/db_xref="GI:36786553"

/translation="MMSERKNEIFTHPEYDELGLPYVVPNARIEENVATCCKYAT
KVLPIVLPQWGSNLSKSTEGSWRLNRLISFDVGMWCSGASRYKTLIDNKTIVD
DSGDIPDHTKXKFKQCKQKQNGEYAHNSYGFPLMWQAVLADDERLAECYLAEGE
KTLRODMVMSLNAEWGEPLTEAEVDHSDYDLYPVHVGWMLQSNESASRLBOY
DKYLAFYKRCATKRVILVTHSMGVLVAYSELNLRGKIGIYGVWVAPGAPPTY
KMKNGEDVATGVLVVGDEDMTPVAOSRGLLPGREYRGMLHINDGWTKPLP
KSPVREIYVZEDRWMLCTETPFLINDKDKKDKGASMSYKILRNTPYPIEILTS
KHAETVAYGSEKRLSYGVISWQASQDYNNKTEDYSGLTFDPTIYDYNLEIGAT
RMVQFSVGSFQDIAKTKPLAPPKPEGGVIVQSGRIAAARYLSLATEVDHGGAY
KQNETELETFAFLFTLRISIVKNOVAKGEN"

gene

complement (1018. .10762)

/locus_tag="plu3237"

complement (8237. .10762)

/locus_tag="plu3237"

/function="Unknown"

/note="unnamed protein product; Similar to YAG related
protein"

/codon_start=1

/transl_table=11

/protein_id="CAE15611.1"

/db_xref="GI:36786554"

/translation="MTNLTPMIIIFDHSRYRLKVMKNAKADLVATFGQALSGPRRYA

```
IEFTSAKEGIEPAOMIMOOASFTLTSPAINPGIRWMPYVEPEPLRSVGVISGICLTS
TSDRSRYEVTLTEPLALASHOVAIYOKLSVPEITVKKILREHGFPGOPFLESLAY
SYPRERQVWQYGEDDLRFQRLALAVGIVTKTADRLKIDIVEFYDORQOPWVYL
PARNBSGHWGDEDAVWGMETAIYQVEGKIATRDYDCHARPPELANTBADYTRDRI
TYGEAYHVDYDRIIPGDRYAYQPETESGVFARLHRLYRNLIRLQWMTSATLVPG
OELIKVGDAPETFRKGAITQITNSARSDSEFMTFTAIPISEIVCFRPERPRKVA
GITPAVSTKVDYDIDKDLGVFVFNFERKMPQGESLWRLARPYAGENVGF
HMPLEGTETVALAEFGGDDPRPIYALAHDSVHPDVMNLXNYKRVLRTPANKRLMD
DGRPHRITLSTEGEGKQSLNIGHVDSORPHDPDGGEGELRTDMDKAIAGKGLFTI
SHDKORACGEVILBEMBALNQCQAOITLFEALYGAENAKATLADLOCKKALISTLA
ELKKSALISAPBGIAQTAKSLQDLAAGSNITVSGKSDFSVLKFTVLAAGERISLFI
ACKLIGTLFASRGRVDIEAGDGMELALAKDITVSHSEKVIISAQKOBILVSGGGYI
RIONGVCEGAPNHIIORAVWQKESGVSQTVQWQCYANVIAVPKAVRAYKXISPLD
ROMQFHADGVAQALSTIONGKTPLOKQGVIEISQLKIDEE"
complement (8242..8247)
/locus_tag="Plu3235"
complement (8242..8247)
/locus_tag="Plu3235"
complement (8242..8247)
/locus_tag="Plu3235"
complement (10767..10772)
/locus_tag="Plu3237"
complement (11601..14951)
/locus_tag="Plu3238"
complement (11601..14951)
/locus_tag="Plu3238"
/function="Unknown"
/Note="unnamed protein product; Similar to putative
membrane protein of Y. pestis"
/codon_start=1
/transl_table=11
/protein_id="CAFI5612.1"
/db_xref="GI:36786555"
/translation="VVKTLILIGITLLIVSAGLGMWVGHRTTDAETENHILMVC
GIAMVLTIVVSMIGITVTLFVNKVTDVKSDTRPLPASEPDYGDGDRRLVYGR
FMWYKRVILVGEVEVQVBAIAGLTTMOEGYRLLMGSGIQAEPNQTQMLNR
LRVYRPLNIVMLTENQSAOPVMDKALRMLOKQAOIRMOAPVLMVOCVSTSOE
GRITQVGGFFPFCRCPDIAVQLOTLVNLREMGQGLLETTAPLFCRISANIKQO
GIAHQVQVITPMLAEYGVYSURGLMFSPLKQAQITRLQITLPTANQGVLEGRCF
```

Query Match 84.0%; Score 16.8; DB 1; Length 348505;
Best Local Similarity 90.0%; Pred. No. 3e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AGTCGCAACAGTACATGG 20
DB 111360 AGACGCAACAGTACAAAG 111379

RESULT 42
TAE566651 925 bp DNA linear PLN 16-MAR-2004
LOCUS Triticum aestivum subsp. aestivum partial Glu-A1-2 gene for
DEFINITION high-molecular-weight glutenin subunit, exon 1, specimen voucher
Gatersleben TRI 10914.
AJ566651
ACCESSION AJ566651.1 GI:32328626
VERSION Glu-A1-2 gene; high-molecular-weight glutenin subunit.
KEYWORDS Triticum aestivum (bread wheat)
SOURCE Triticum aestivum
ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Poideae; Triticaceae; Triticum.

REFERENCE 1

Athlet R.H., Jacomet S. and SchLundbaum A.

About the origin of European spelt (Triticum spelta L.): allelic
differentiation of the HMW Glutenin B1-1 and A1-2 subunit genes
Theor. Appl. Genet. 108 (2), 360-367 (2004)

JOURNAL 14564390

PUBMED 2 (bases 1 to 925)

REFERENCE Blatter R.H.E.

AUTHORS Direct Submission

TITLE Submitted (18-MAY-2003) Blatter R.H.E., Genetics MPI for

Evolutionary Anthropology, Deutscher Platz, 04105 Leipzig, GERMANY
Location/Qualifiers

source 1..925
/organism="Triticum aestivum"
/mol_type="genomic DNA"
/sub_species="aestivum"
/specimen_voucher="Gatersleben TRI 10914 (IPK)"
/db_xref="taxon:4565"
/tissue_type="leaf"
/country="Nepal"
/note="cultivar-group: villosum"
423..925
/gene="Glu-A1-2"
423..429
/gene="Glu-A1-2"
514..>925
/gene="Glu-A1-2"
/codon_start=1
/product="high-molecular-weight glutenin subunit"
/protein_id="CAE00513.1"
/db_xref="GI:32328627"
/db_xref="TrEMBL:Q7X6P9"
/translation="MAKRLVFAVTVIGLVSLTVABGEASKQDQCRRLQESSLEACR
LVVDQLAGRLPMSTGLQMRCCOQRLDISAKRPVALGVARQYGTAVPPKGSFVH
RFTQQLQQLQGIQGTSSQTVQGYPSVISPQGG"
<514..>925
/gene="Glu-A1-2"
/number=1

Query Match 82.0%; Score 16.4; DB 8; Length 925;
Best Local Similarity 94.4%; Pred. No. 6.6e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 TCGCAACAGTACATGG 20
DB 874 TCGCAACAGTACAAAG 891

RESULT 43
CQ574602 950 bp DNA linear PAT 02-FEB-2004
LOCUS CQ574602
DEFINITION Sequence 2360 from Patent WO0171042.
ACCESSION CQ574602
VERSION CQ574602.1 GI:41638381
KEYWORDS Drosophila sp.
SOURCE Drosophila sp.
ORGANISM Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.

REFERENCE 1
Venter, J.C., Adams, M., Li, P.W. and Myers, E.W.
Detection kits, such as nucleic acid arrays, for detecting the
expression of 10,000 or more Drosophila genes and uses therefor
Patent: WO 0171042-A 2360 27-SEP-2001;
PE Corporation (NY) (US)

LOCATION/Qualifiers

1..950
/organism="Drosophila sp."
/mol_type="unassigned DNA"
/db_xref="taxon:7242"

ORIGIN

Query Match 82.0%; Score 16.4; DB 6; Length 950;
Best Local Similarity 94.4%; Pred. No. 6.6e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 TCGCAACAGTACATGG 20
DB 634 TCGCAACAGTACATGG 651

RESULT 44
AY060813

LOCUS AY060813 1143 bp mRNA linear INV 08-NOV-2001
 DEFINITION Drosophila melanogaster GH28342 full length cDNA.
 ACCESSION AY060813
 VERSION AY060813.1 GI:16768283
 KEYWORDS F11 cDNA.
 SOURCE Drosophila melanogaster (fruit fly)
 ORGANISM Drosophila melanogaster
 Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
 1 (bases 1 to 1143)
 AUTHORS Stapleton, M., Broksstein, P., Hong, L., Agdayani, A., Carlson, J., Champe, M., Chavez, C., Dorsett, V., Farfan, D., Fitse, E., Georges, R., Gonzalez, M., Guarin, H., Li, P., Liao, G., Miranda, A., Mungall, C. J., Nuno, J., Pacle, J., Paragas, V., Park, S., Phouanavong, S., Wan, K., Yu, C., Lewis, S. E., Rubin, G. M. and Celisner, S.
 TITLE Submitted (30-OCT-2001) Berkeley Drosophila Genome Project, Lawrence Berkeley National Laboratory, One Cyclotron Road, Berkeley, CA 94720, USA
 JOURNAL
 COMMENT
 Sequence submitted by:
 Berkeley Drosophila Genome Project
 Lawrence Berkeley National Laboratory
 Berkeley, CA 94720
 This clone was sequenced as part of a high-throughput process to sequence clones from Drosophila Gene Collection 1 (Rubin et al., Science 2000). The sequence has been subjected to integrity checks for sequence accuracy, presence of a polyA tail and contiguity within 100 kb in the genome. Thus we believe the sequence to reflect accurately this particular cDNA clone. However, there are artifacts associated with the generation of cDNA clones that may have not been detected in our initial analyses such as internal priming, priming from contaminating genomic DNA, retained introns due to reverse transcription errors that result in single base changes. For further information about this sequence, including its location and relationship to other sequences, please visit our web site (<http://fruitfly.berkeley.edu>) or send email to conae@fruitfly.berkeley.edu.
 Location/Qualifiers
 1..1143
 /organism="Drosophila melanogaster"
 /mol_type="mRNA"
 /strain="y; cn bw sp"
 /db_xref="taxon:7227"
 /map="46A3-46A3"
 1..1143
 /gene="CG1773"
 /note="alignment with genomic scaffold AE003832"
 /db_xref="FLYBASE:FBgn0033439"
 9..962
 /gene="CG1773"
 /note="Longest ORF"
 /codon_start=1
 /product="GH28342p"
 /protein_id="AAL28361.1"
 /db_xref="GI:16768284"
 /db_xref="FLYBASE:FBgn0033439"
 /translation="MNSPLGITALIWGILICSCPPSSQAGREDWPHETLAVEDTQDGVLSNLPAPRLRRITGKSSLSLQPMARFHIISGDIEMCRGSSLSLSELVITAAHCFMCPKREIRWIGELDISSTDCVYNVQVCAALPYEETIDKMLHEFNLFYGDIALIKLNKRVFMDHRIPLCLPTLELLAFITLQGSYVAWGMRTESRRFANSTMEVHINTKCTGDRDTSFLCANGDYDTCTGSGSLPWKTLTFKARTVFGVYSTGSGCAGQAYVMDVPTVYFMIILAKKAEISDPKSLHR"
 ORIGIN
 Query Match 82.0%; Score 16.4; DB 3; Length 1143;
 Best Local Similarity 94.4%; Pred. No. 6.6e+02;
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 3 TCGCAACAGTACATG 20
 |||||
 658 TCGCCACAGTACATG 675

RESULT 45
 LOCUS COS94254/c 2901 bp DNA linear PAT 02-FEB-2004
 DEFINITION Sequence 22012 from Patent WO0171042.
 ACCESSION COS94254
 VERSION COS94254.1 GI:1651482
 KEYWORDS
 SOURCE
 ORGANISM Drosophila sp.
 Drosophila sp.
 Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
 1
 AUTHORS Venter, J. C., Adams, M., Li, P. W. and Myers, E. W.
 TITLE Detection kits, such as nucleic acid arrays, for detecting the expression of 10,000 or more Drosophila genes and uses thereof
 JOURNAL Patent: WO 0171042-A 22012 27-SEP-2001; PE Corporation (NY) (US)
 FEATURES
 source
 1..2901
 /organism="Drosophila sp."
 /mol_type="unassigned DNA"
 /db_xref="taxon:7242"
 ORIGIN
 Query Match 82.0%; Score 16.4; DB 6; Length 2901;
 Best Local Similarity 94.4%; Pred. No. 6.3e+02;
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 3 TCGCAACAGTACATG 20
 |||||
 2408 TCGCCACAGTACATG 2391

RESULT 46
 LOCUS COS74601/c 3165 bp DNA linear PAT 02-FEB-2004
 DEFINITION Sequence 2359 from Patent WO0171042.
 ACCESSION COS74601
 VERSION COS74601.1 GI:1638380
 KEYWORDS
 SOURCE
 ORGANISM Drosophila sp.
 Drosophila sp.
 Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
 1
 AUTHORS Venter, J. C., Adams, M., Li, P. W. and Myers, E. W.
 TITLE Detection kits, such as nucleic acid arrays, for detecting the expression of 10,000 or more Drosophila genes and uses thereof
 JOURNAL Patent: WO 0171042-A 2359 27-SEP-2001; PE Corporation (NY) (US)
 FEATURES
 source
 1..3165
 /organism="Drosophila sp."
 /mol_type="unassigned DNA"
 /db_xref="taxon:7242"
 ORIGIN
 Query Match 82.0%; Score 16.4; DB 6; Length 3165;
 Best Local Similarity 94.4%; Pred. No. 6.3e+02;
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 3 TCGCAACAGTACATG 20
 |||||
 1317 TCGCCACAGTACATG 1300

RESULT 47
 LOCUS AC017970 89765 bp DNA linear HTG 09-DEC-1999
 AC017970

DEFINITION Drosophila melanogaster, *** SEQUENCING IN PROGRESS ***.
ACCESSION AC01970
VERSION AC01970.1 GI:6553220
KEYWORDS HTG; HTGS_PHASE2.
SOURCE Drosophila melanogaster (fruit fly)
ORGANISM Drosophila melanogaster
Bukayocra; Metazoa; Arthropoda; Hexapoda; Insecta; Plearyocra;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 89765)
AUTHORS Adams, M. and Venter, J.C.
TITLE Direct Submission
JOURNAL Submitted (09-DEC-1999) Celera Genomics, 45 West Gude Drive,
Rockville, MD, USA
COMMENT This sequence was identified as CM:1021823 by the submitter.
For more information on this record e-mail to fly@celera.com.
* NOTE: This is a 'working draft' sequence.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
FEATURES
source
1..89765
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
ORIGIN
Query Match 82.0%; Score 16.4; DB 2; Length 89765;
Best Local Similarity 94.4%; Pred. No. 5.3e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 3 TCGCAACAGTACATCG 20
|||||||
Db 81121 TCGCAACAGTACATCG 81138
RESULT 48
AL390057 100578 bp DNA linear PRI 25-MAR-2001
LOCUS Human DNA sequence from clone RP11-26118 on chromosome 6 Contains
DEFINITION STS and GSSs, complete sequence.
ACCESSION AL390057
VERSION AL390057.12 GI:11322100
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 100578)
REFERENCE
AUTHORS Williams, S.
TITLE Direct Submission
JOURNAL Submitted (23-MAR-2001) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
On Nov 23, 2000 this sequence version replaced gi:1113956.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
The following abbreviations are used to associate primary accession
numbers given in the feature table with their source databases:
Em: EMBL; Sw: SWISSPROT; Tr: TREMBL; Wp: WORMPEP; Information
on the WORMPEP database can be found at
http://www.sanger.ac.uk/Projects/C.elegans/wormpep This sequence
was generated from part of bacterial clone contigs of human
chromosome 6, constructed by the Sanger Centre Chromosome 6 Mapping
Group. Further information can be found at
http://www.sanger.ac.uk/HGP/chr6
IMPORTANT: This sequence is not the entire insert of clone
RP11-26118 it may be shorter because we sequence overlapping
sections only once, except for a 100 base overlap.

FEATURES
source
The true right end of clone RP11-26118 is at 100578 in this
sequence. The true left end of clone RP11-300612 is at 62151 in this
sequence. The true right end of clone RP11-67K21 is at 100 in this
sequence. This sequence was finished as follows unless otherwise
noted: all regions were either double-stranded or sequenced with an
alternate chemistry or covered by high quality data (i.e., phred
quality >= 30); an attempt was made to resolve all sequencing
problems, such as compressions and repeats; all regions were
covered by at least one plasmid subclone or more than one M13
subclone; and the assembly was confirmed by restriction digest.
RP11-26118 is from the library RPCI-11.1 constructed by the group
of Pieter de Jong. For further details see
http://www.chori.org/Dacpac/home.htm
VECTOR: PBAC3.6.
Location/Qualifiers
1..100578
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="6"
/clone="RP11-26118"
/clone_id="RPCI-11.1"
101..740
/note="match: GSS: Em:AQ302222"
155..1072
/note="L1PA3 repeat: matches 5230. .6146 of consensus"
1078..1291
/note="L1PA3 repeat: matches 5022. .5235 of consensus"
/note="match: GSS: Em:AQ043864"
6359..6945
/note="match: GSS: Em:AQ296519"
9258..10681
/note="L1PA7 repeat: matches 3942. .5399 of consensus"
10984..11725
/note="L1PA7 repeat: matches 5399. .6124 of consensus"
/note="match: GSS: Em:AQ618053"
12918..13165
/note="match: GSS: Em:AQ054828"
12990..13356
/note="match: GSS: Em:AQ588764"
13504..13645
/note="L1Mcb repeat: matches 789. .930 of consensus"
13901..13940
/note="L1Mcb repeat: matches 828. .930 of consensus"
14079..14422
/note="L1Mcb repeat: matches 143. .443 of consensus"
15727..15784
/note="L1Mcb repeat: matches 6239. .6296 of consensus"
16035..16272
/note="match: GSS: Em:AQ042985"
20715..20810
/note="L1Mcb repeat: matches 6098. .6304 of consensus"
22827..23016
/note="L1Mcb repeat: matches 6098. .6304 of consensus"
/note="match: GSS: Em:AQ486821"
26590..26801
/note="match: GSS: Em:AQ486821"
30142..30175
/note="L1Mcb repeat: matches 3. .1579 of consensus"
30620..30675
/note="L1Mcb repeat: matches 3. .1579 of consensus"
30884..32217
/note="L1Mcb repeat: matches 2. .2418 of consensus"
34613..34688
/note="L1Mcb repeat: matches 2. .2418 of consensus"
34773..36379
/note="L1Mcb repeat: matches 2. .2418 of consensus"
36240..36273
/note="L1Mcb repeat: matches 4514. .6146 of consensus"
36380..36413
/note="Sequence confirmed by AC016105 sequenced by W1BR."

```

repeat_region /note="17 copies 2 mer aa 82% conserved"
               /note="LIPB5 repeat: matches 5770. .6140 of consensus"
repeat_region 38583.38818
               /note="LIM4 repeat: matches 5931. .6164 of consensus"
misc_feature 41526.41977
               /note="match: GSS: Em:AQ126191"
misc_feature 41540.42000
               /note="match: GSS: Em:AQ613376"
misc_feature 41548.42098
               /note="match: GSS: Em:AQ378516"
misc_feature 41552.42000
               /note="match: GSS: Em:AQ703614"
repeat_region 42853.43035
               /note="LIPB1 repeat: matches 5962. .6144 of consensus"
repeat_region 43625.43704
               /note="40 copies 2 mer ag 73% conserved"
repeat_region 47701.48062
               /note="LIPB2 repeat: matches 3. .374 of consensus"
misc_feature complement(48868. .49490)
               /note="match: GSS: Em:AQ109429"
misc_feature 49283.49723
               /note="match: GSS: Em:AQ803593"
repeat_region 49524.49555
               /note="16 copies 2 mer tg 90% conserved"
repeat_region 50381.50723
               /note="TIGER1 repeat: matches 2045. .2415 of consensus"
repeat_region 51284.51307
               /note="12 copies 2 mer aa 100% conserved"
repeat_region 56705.57064
               /note="THIC repeat: matches 1. .371 of consensus"
repeat_region 57065.58634
               /note="THIC-internal repeat: matches 1. .1580 of consensus"
repeat_region 58637.58939
               /note="THIC repeat: matches 5. .371 of consensus"
misc_feature 60208.60748
               /note="match: GSS: Em:AQ389684"
repeat_region 60770.60825
               /note="28 copies 2 mer aa 71% conserved"
misc_feature complement(60839. .61518)
               /note="match: GSS: Em:AQ308800"
repeat_region 62037.62244
               /note="LIM4 repeat: matches 6102. .6295 of consensus"
repeat_region 62539.62734
               /note="LIM4 repeat: matches 5917. .6115 of consensus"
repeat_region 62854.65962
               /note="LIM4 repeat: matches 2760. .5917 of consensus"
repeat_region 65369.68471
               /note="LIPB3 repeat: matches 3648. .6146 of consensus"
repeat_region 68476.72115
               /note="LIPB3 repeat: matches 11. .3642 of consensus"
repeat_region 72327.73398
               /note="LIPB3 repeat: matches 4581. .6148 of consensus"
repeat_region 75859.76732
               /note="MER1D repeat: matches 1. .897 of consensus"
repeat_region 76897.77556
               /note="PAB1A repeat: matches 1. .660 of consensus"
repeat_region 79727.79864
               /note="68 copies 2 mer aa 68% conserved"
misc_feature 83250.83780
               /note="match: GSS: Em:AQ566452"
misc_feature 83271.83732
               /note="match: GSS: Em:AQ420577"
misc_feature complement(85140. .85328)
               /note="match: STS: Em:HSPB31F2"
misc_feature complement(86967. .89333)
               /note="match: GSS: Em:AQ242452"
repeat_region 89334.89573
               /note="MER30 repeat: matches 1. .230 of consensus"
misc_feature 92101.92396
               /note="match: GSS: Em:AQ66482"
repeat_region 93663.93702

```

```

repeat_region /note="20 copies 2 mer ta 82% conserved"
               95778.97415
               /note="LIPB2 repeat: matches 4513. .6155 of consensus"
repeat_region 97414.98915
               /note="LIPB2 repeat: matches 2970. .4484 of consensus"
repeat_region 99680.99757
               /note="139 copies 2 mer ta 71% conserved"
misc_feature complement(99955. .100244)
               /note="match: GSS: Em:AQ388863"
misc_feature complement(99960. .100368)
               /note="match: GSS: Em:AQ388862"
misc_feature match: STS: Em:G59389"
               complement(100039. .100572)
               /note="match: GSS: Em:AQ416636"
               complement(100159. .100554)
               /note="match: GSS: Em:AQ003409
               match: STS: Em:G51488"

ORIGIN
Query Match      82.0%; Score 16.4; DB 9; Length 100578;
Best Local Similarity 94.4%; Pred. No. 5.3e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Cy      2 GTGGCAACAGTCAATG 19
Db      80450 GTACCAACAGTCAATG 80467

RESULT 49
BX649472      108623 bp      DNA      linear      VRT 17-FEB-2004
LOCUS      Zebrafish DNA sequence from clone CH211-214K in linkage group 20,
DEFINITION      complete sequence.
ACCESSION      BX649472
VERSION      BX649472.15 GI:42592727
KEYWORDS      HTG.
SOURCE      Danio rerio (zebrafish)
ORGANISM      Danio rerio
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
              Cypriniformes; Cyprinidae; Danio.
              1 (bases 1 to 108623)
REFERENCE      Dyer, L.
              Direct Submission
              Submitted (14-FEB-2004) Wellcome Trust Sanger Institute, Hinxton,
              Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
              zfish-help@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
              On Feb 17, 2004 this sequence version replaced GI:41222936.
COMMENT      ----- Genome Center
              Center: Wellcome Trust Sanger Institute
              Center code: SC
              Web site: http://www.sanger.ac.uk
              Contact: zfish-help@sanger.ac.uk

```

During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YAC.

The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em, EMBL; Sw, SWISSPROT; Tr, TrEMBL; Wp, WormPep; Infomation on the WormPep database can be found at http://www.sanger.ac.uk/Projects/C_elegans/wormpep Clone-derived

zebrafish pUC subclones occasionally display inconsistency over the length of mononucleotide A/T runs and conserved TA repeats. Where this is found the longest good quality representation will be submitted.

Repeat names beginning 'Dr' were identified by the Recon repeat discovery system (Zhirong Bao and Sean Eddy, submitted), and those beginning 'drr' were identified by Rick Waterman (Stephen Johnson lab, WashU). For further information see http://www.sanger.ac.uk/projects/D_rexio/fishmask.shtml CH211-214K9 is from a CHOR1-211 BAC library

FEATURES

source

location/Qualifiers
1..108623
/organism="Danio rerio"
/mol_type="genomic DNA"
/db_xref="taxon:7955"
/clone="CH211-214K9"
/clone_lib="CHOR1-211"

ORIGIN

Query Match 82.0%; Score 16.4; DB 5; Length 108623;
Best Local Similarity 94.4%; Pred. No. 5.3e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AGTCGCAACAGTACAT 18
|||
Db 46618 AGTTCGCAACAGTACAT 46635

RESULT 50
AL365229/c 121478 bp DNA linear PRI 14-FEB-2001
LOCUS Human DNA sequence from clone RP11-723K16 on chromosome 20 Contains
DEFINITION ESTs, STSs, GSSs and two Cpg islands. Contains the 5' end of the
CDH4 gene for retinal cadherin 4 type 1 (R-cadherin), complete
sequence.

ACCESSION AL365229 GI:10443454
VERSION AL365229.13 HTG; cadherin; CDH4; Cpg island.
KEYWORDS Homo sapiens (human)
SOURCE
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
TITLE 1 (bases 1 to 121478)
JOURNAL Wilson, S.

COMMENT Submitted (09-FEB-2001) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk Clone
requests: clonerequests@sanger.ac.uk

COMMENT

On Oct 1, 2000 this sequence version replaced gi:10186767.

During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases:

Em: EMBL; Sw: SWISSPROT; Tr: TREMBL; Wp: WORMPEP; Information on the WORMPEP database can be found at

http://www.sanger.ac.uk/projects/C_elegans/wormpep This sequence was generated from part of bacterial clone contigs of human 20

chromosome 20, constructed by the Sanger Centre Chromosome 20 Mapping Group. Further information can be found at

<http://www.sanger.ac.uk/HGP/Chr20>

IMPORTANT: This sequence is not the entire insert of clone RP11-723K16. It may be shorter because we sequence overlapping sections only once, except for a 100 base overlap.

The true right end of clone RP11-723K16 is at 121478 in this sequence. The true right end of clone RPS-1040G13 is at 100 in this sequence. RP11-723K16 is from the library RPI-11.3 constructed by

the group of Pieter de Jong. For further details see <http://www.chori.org/bacpac/home.htm>

FEATURES

source

VECTOR: pBAC3.6
This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest.

location/Qualifiers

1..121478
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="20"
/clone="RP11-723K16"
/clone_lib="RPI-11.3"
complement(1..76)
/note="match: STS: Em:HS040G13S"
1480..1588
/note="12 repeat: matches 2606..2716 of consensus"
1580..1920
/note="match: GSS: Em:BL7526"
1612..1738
/note="match: GSS: Em:BL5710"
1683..2095
/note="MER66A repeat: matches 27..471 of consensus"
2356..2603
/note="ULM4 repeat: matches 6694..6947 of consensus"
3193..3283
/note="WIR repeat: matches 118..221 of consensus"
3619..3865
/note="WLR1J repeat: matches 293..547 of consensus"
3866..4159
/note="ALUSx repeat: matches 1..293 of consensus"
4160..4314
/note="WLR1H repeat: matches 1..293 of consensus"
4894..4929
/note="9 copies 4 mer ttta 83% conserved"
7552..7681
/note="MER5B repeat: matches 38..166 of consensus"
7906..8195
/note="ALUSx repeat: matches 1..296 of consensus"
8365..8779
/note="ULM3C repeat: matches 1245..1669 of consensus"
9548..9583
/note="18 copies 2 mer tg 88% conserved"
9554..9585
/note="8 copies 4 mer tgtg 96% conserved"
10035..11189
/note="ULM4 repeat: matches 6696..7851 of consensus"
11976..12219
/note="WLR1J repeat: matches 275..514 of consensus"
12442..12502
/note="ULM3 repeat: matches 6042..6102 of consensus"
12552..12746
/note="3 copies 65 mer 81% conserved"
12356..12609
/note="9 copies 6 mer tcgag 77% conserved"
12615..12680
/note="11 copies 6 mer tgltag 83% conserved"
12685..13044
/note="12 copies 30 mer 77% conserved"
12700..13053
/note="59 copies 6 mer tagtg 78% conserved"
12856..12859
/note="Single clone region"
13050..13289
/note="16 copies 15 mer 64% conserved"
13053..13256
/note="68 copies 3 mer gtc 65% conserved"
13072..13311
/note="8 copies 30 mer 87% conserved"
13098..13299


```
repeat_region /note="101 copies 2 mer gt 55% conserved"
13253. .13382 /note="2 copies 65 mer 83% conserved"
repeat_region 13255. .13320
/note="11 copies 6 mer tgtagg 75% conserved"
repeat_region 13316. .13687
/note="124 copies 3 mer gag 55% conserved"
repeat_region 13321. .13686
/note="61 copies 6 mer gtaggt 78% conserved"
repeat_region 13687. .13728
/note="7 copies 6 mer gtgcag 81% conserved"
repeat_region 14398. .14489
/note="MER53 repeat: matches 1. .95 of consensus"
repeat_region 14502. .14794
/note="AluX repeat: matches 1. .293 of consensus"
repeat_region 15370. .15476
/note="L2 repeat: matches 2625. .2742 of consensus"
repeat_region 16777. .17289
/note="MLTID repeat: matches 2. .505 of consensus"
repeat_region 17603. .17729
/note="MIR repeat: matches 28. .153 of consensus"
repeat_region 17810. .17887
/note="MAD1 repeat: matches 1. .80 of consensus"
repeat_region 18309. .18503
/note="MER3 repeat: matches 11. .209 of consensus"
repeat_region 19972. .20149
/note="MER5B repeat: matches 1. .177 of consensus"
repeat_region 20569. .20977
/note="LIPB1 repeat: matches 5747. .6152 of consensus"
repeat_region 21167. .21517
/note="LIMB7 repeat: matches 5769. .6130 of consensus"
repeat_region 22885. .23003
/note="MLTID repeat: matches 1. .121 of consensus"
repeat_region 23015. .23157
/note="MLTID repeat: matches 261. .452 of consensus"
repeat_region 23618. .23673
/note="2 copies 28 mer 92% conserved"
repeat_region 23619. .23674
/note="8 copies 7 mer ctgggct 92% conserved"
repeat_region 24512. .24579
/note="17 copies 4 mer atat 100% conserved"
repeat_region 24632. .24934
/note="AluX repeat: matches 1. .306 of consensus"
misc_feature complement(24989. .25460)
/note="match: GSS: Em:AQ721413"
repeat_region 25628. .26093
/note="LIMC/D repeat: matches 5331. .5763 of consensus"
repeat_region 26125. .26179
/note="LIMB1 repeat: matches 6100. .6156 of consensus"
repeat_region 26839. .27148
/note="MLTIA2 repeat: matches 1. .316 of consensus"
repeat_region 27569. .27718
/note="3 copies 50 mer 78% conserved"
repeat_region 27570. .27719
/note="5 copies 30 mer 74% conserved"
repeat_region 27597. .27716
/note="8 copies 15 mer 65% conserved"
repeat_region 27908. .28211
/note="LIM3 repeat: matches 216. .184 of consensus"
repeat_region 28586. .28685
/note="25 copies 4 mer tgtg 74% conserved"
repeat_region 28591. .28678
/note="4 copies 22 mer 76% conserved"
repeat_region 28592. .28675
/note="3 copies 28 mer 81% conserved"
repeat_region 28681. .28830
/note="5 copies 30 mer 73% conserved"
repeat_region 28687. .28818
/note="22 copies 6 mer tgtgtg 77% conserved"
repeat_region 28688. .28819
/note="3 copies 44 mer 83% conserved"
repeat_region 28693. .28820
/note="32 copies 4 mer tgtg 78% conserved"
```

```
repeat_region 28703. .28812 /note="5 copies 22 mer 80% conserved"
misc_feature complement(30513. .30994)
/note="match: GSS: Em:AQ454650"
repeat_region 30702. .30783 /note="2 copies 41 mer 87% conserved"
repeat_region 31320. .31610 /note="AluX repeat: matches 4. .294 of consensus"
repeat_region 32334. .32778 /note="L1 repeat: matches 2672. .3123 of consensus"
repeat_region 32763. .33175 /note="LIMB1 repeat: matches 5675. .6089 of consensus"
```

```
Query Match 82.0%; Score 16.4; DB 9; Length 121478;
Best Local Similarity 94.4%; Pred. No. 5.3e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 AGTGCACAAGTACAT 18
Db 5869 AGTGCACAAGTACAT 5852
```

Search completed: December 3, 2004, 03:07:41
Job time : 704.579 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Comphen Ltd.

CM nucleic - nucleic search, using sw model

Run on: December 3, 2004, 02:05:35 ; Search time 1585.42 Seconds
(without alignments)
62.385 Million cell updates/sec

Title: US-10-050-189a-6
Perfect score: 18
Sequence: 1 gagacaacaagatcgc 18

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3694831 segs, 274740616 residues
Total number of hits satisfying chosen parameters: 7389662

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 1000 summaries

Database : Published Applications NA.*
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq.*
10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq.*
11: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq.*
12: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq.*
13: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq.*
14: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq.*
15: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq.*
16: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq.*
17: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq.*
18: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq.*
19: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq.*
20: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*
21: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	18	100.0	18	US-10-050-189a-6	Sequence 6, Appl1
2	16	88.9	2465	US-10-027-632-111967	Sequence 111967,
3	16	88.9	15	US-10-027-632-111967	Sequence 111967,
4	16	88.9	66479	US-10-041-856-1	Sequence 1, Appl1
5	15.4	85.6	409	US-09-833-381-908	Sequence 908, Appl1
6	15.4	85.6	482	US-09-878-178-1155	Sequence 1155, Ap
7	15.4	85.6	482	US-10-046-935-1155	Sequence 1155, Ap
8	15.4	85.6	482	US-10-146-502-1155	Sequence 1155, Ap
9	15.4	85.6	513	US-10-282-122A-20772	Sequence 20772, A
10	15.4	85.6	577	US-10-027-632-277903	Sequence 277903, A
11	15.4	85.6	577	US-10-027-632-277903	Sequence 277903, A
12	15.4	85.6	582	US-10-027-632-277719	Sequence 277719, A

13	15.4	85.6	582	US-10-027-632-277719	Sequence 277719,
14	15.4	85.6	582	US-10-767-701-3256	Sequence 3256, Ap
15	15.4	85.6	775	US-10-437-963-32919	Sequence 32919, A
16	15.4	85.6	1470	US-09-815-242-6118	Sequence 6118, Ap
17	15.4	85.6	1470	US-10-369-493-24585	Sequence 24585, A
18	15.4	85.6	1518	US-10-369-493-32483	Sequence 32483, A
19	15.4	85.6	1950	US-10-369-493-45765	Sequence 45765, A
20	15.4	85.6	2186	US-10-169-257-8	Sequence 8, Appl1
21	15.4	85.6	3178	US-09-308-207-7	Sequence 7, Appl1
22	15.4	85.6	5101	US-10-093-463-135	Sequence 135, App
23	15.4	85.6	5719	US-10-367-094-100	Sequence 100, App
24	15.4	85.6	7043	US-09-070-927A-288	Sequence 288, App
25	15.4	85.6	142976	US-10-367-094-99	Sequence 99, Appl
26	15	83.3	378	US-10-437-963-86110	Sequence 86110, A
27	15	83.3	485	US-10-137-036-4	Sequence 4, Appl1
28	15	83.3	737	US-09-770-149-77	Sequence 77, Appl
29	15	83.3	1169	US-10-767-701-11437	Sequence 11437, A
30	15	83.3	1455	US-10-032-585-6133	Sequence 6133, Ap
31	14.8	82.2	209	US-10-425-115-87011	Sequence 87011, A
32	14.8	82.2	352	US-09-867-550-1599	Sequence 1599, A
33	14.8	82.2	416	US-10-425-115-91075	Sequence 91075, A
34	14.8	82.2	438	US-10-282-122A-33603	Sequence 33603, A
35	14.8	82.2	452	US-10-425-114-3291	Sequence 3291, Ap
36	14.8	82.2	462	US-10-767-701-26974	Sequence 26974, A
37	14.8	82.2	465	US-09-918-995-1845	Sequence 1845, Ap
38	14.8	82.2	467	US-10-260-238-4785	Sequence 4785, Ap
39	14.8	82.2	470	US-10-425-115-81576	Sequence 81576, A
40	14.8	82.2	473	US-10-027-632-68850	Sequence 68850, A
41	14.8	82.2	473	US-10-027-632-68851	Sequence 68851, A
42	14.8	82.2	473	US-10-027-632-68850	Sequence 68850, A
43	14.8	82.2	473	US-10-027-632-68851	Sequence 68851, A
44	14.8	82.2	480	US-10-424-599-16732	Sequence 16732, A
45	14.8	82.2	480	US-10-424-599-85425	Sequence 85425, A
46	14.8	82.2	585	US-10-027-632-217259	Sequence 217259, A
47	14.8	82.2	585	US-10-027-632-217260	Sequence 217260, A
48	14.8	82.2	585	US-10-027-632-217259	Sequence 217259, A
49	14.8	82.2	585	US-10-027-632-217260	Sequence 217260, A
50	14.8	82.2	666	US-10-425-114-5996	Sequence 5996, Ap
51	14.8	82.2	671	US-10-027-632-300982	Sequence 300982, A
52	14.8	82.2	671	US-10-027-632-300982	Sequence 300982, A
53	14.8	82.2	676	US-10-313-542-15	Sequence 15, Appl1
54	14.8	82.2	689	US-10-425-114-16588	Sequence 16588, A
55	14.8	82.2	699	US-10-425-114-7195	Sequence 7195, Ap
56	14.8	82.2	701	US-10-425-115-115987	Sequence 115987, A
57	14.8	82.2	723	US-10-437-963-16634	Sequence 16634, A
58	14.8	82.2	727	US-10-425-114-23906	Sequence 23906, A
59	14.8	82.2	733	US-10-425-114-7249	Sequence 7249, Ap
60	14.8	82.2	736	US-10-425-114-1610	Sequence 1610, Ap
61	14.8	82.2	736	US-10-425-114-22570	Sequence 22570, A
62	14.8	82.2	741	US-10-425-114-18307	Sequence 18307, A
63	14.8	82.2	742	US-10-425-114-25243	Sequence 25243, A
64	14.8	82.2	765	US-10-425-114-1374	Sequence 1374, Ap
65	14.8	82.2	784	US-10-767-795-3053	Sequence 3053, Ap
66	14.8	82.2	795	US-10-653-047-4852	Sequence 4852, Ap
67	14.8	82.2	818	US-10-425-114-21956	Sequence 21956, A
68	14.8	82.2	818	US-10-425-114-21956	Sequence 21956, A
69	14.8	82.2	857	US-10-425-115-5872	Sequence 5872, Ap
70	14.8	82.2	858	US-10-282-122A-41774	Sequence 41774, A
71	14.8	82.2	900	US-10-437-963-73718	Sequence 73718, A
72	14.8	82.2	931	US-10-425-115-114244	Sequence 114244, A
73	14.8	82.2	943	US-10-363-445A-25013	Sequence 25013, A
74	14.8	82.2	943	US-10-363-445A-25014	Sequence 25014, A
75	14.8	82.2	944	US-10-767-701-13612	Sequence 13612, A
76	14.8	82.2	975	US-10-767-701-10956	Sequence 10956, A
77	14.8	82.2	1014	US-10-653-047-575	Sequence 575, App
78	14.8	82.2	1050	US-10-437-963-15052	Sequence 15052, A
79	14.8	82.2	1167	US-10-437-963-96110	Sequence 96110, A
80	14.8	82.2	1212	US-10-250-238-796	Sequence 796, App
81	14.8	82.2	1245	US-10-437-963-70738	Sequence 70738, A
82	14.8	82.2	1315	US-10-369-493-23991	Sequence 23991, A
83	14.8	82.2	1315	US-10-369-493-23991	Sequence 23991, A
84	14.8	82.2	1381	US-10-767-701-13509	Sequence 13509, A
85	14.8	82.2	1383	US-10-282-122A-41827	Sequence 41827, A

C 86	14.8	82.2	1412	17	US-10-437-963-6447	Sequence 6447, App	159	14.4	80.0	1168	18	US-10-425-115-177042	Sequence 177042,
87	14.8	82.2	1511	18	US-10-739-930-2432	Sequence 2432, App	C 150	14.4	80.0	1200	17	US-09-985-153-23	Sequence 23, Appl
88	14.8	82.2	1624	16	US-10-425-114-25941	Sequence 25941, A	161	14.4	80.0	1212	11	US-10-437-963-12980	Sequence 12980, A
89	14.8	82.2	1850	16	US-10-425-114-9032	Sequence 9032, Ap	162	14.4	80.0	1262	17	US-10-437-963-62663	Sequence 62663, A
90	14.8	82.2	1849	16	US-10-424-599-108745	Sequence 108745, A	163	14.4	80.0	1416	16	US-10-425-114-33433	Sequence 33433, A
91	14.8	82.2	1857	13	US-10-027-633-258922	Sequence 258922, A	164	14.4	80.0	1423	9	US-09-764-877-3760	Sequence 3760, Ap
92	14.8	82.2	1867	13	US-10-027-633-258922	Sequence 258922, A	165	14.4	80.0	1423	6	US-10-442-513-3760	Sequence 3760, Ap
93	14.8	82.2	1867	15	US-10-027-633-258922	Sequence 258922, A	166	14.4	80.0	1460	8	US-10-425-114-30890	Sequence 30890, A
94	14.8	82.2	1867	15	US-10-027-633-258922	Sequence 258922, A	C 167	14.4	80.0	1462	16	US-08-781-986A-771	Sequence 771, App
95	14.8	82.2	1923	18	US-10-425-115-149404	Sequence 149404, A	C 168	14.4	80.0	1462	16	US-10-329-62A-771	Sequence 771, App
C 96	14.8	82.2	1954	16	US-10-425-114-26365	Sequence 26365, A	169	14.4	80.0	1662	17	US-10-437-963-62661	Sequence 62661, A
C 97	14.8	82.2	2014	16	US-10-425-114-12937	Sequence 12937, A	170	14.4	80.0	1823	15	US-10-037-270-516	Sequence 516, App
C 98	14.8	82.2	2084	16	US-10-425-114-27735	Sequence 27735, A	171	14.4	80.0	1823	15	US-10-117-722-516	Sequence 516, App
C 99	14.8	82.2	2174	18	US-10-425-115-33959	Sequence 33959, A	172	14.4	80.0	1834	16	US-10-425-114-4429	Sequence 4429, App
C 100	14.8	82.2	2174	18	US-10-425-115-33959	Sequence 33959, A	C 173	14.4	80.0	1890	16	US-10-425-115-16682	Sequence 16682, A
C 101	14.8	82.2	2468	16	US-10-282-122A-26301	Sequence 26301, A	174	14.4	80.0	1906	18	US-10-425-115-71755	Sequence 71755, A
C 102	14.8	82.2	2495	17	US-10-437-963-74259	Sequence 74259, A	175	14.4	80.0	1920	18	US-10-425-115-171486	Sequence 171486, A
C 103	14.8	82.2	2542	18	US-10-425-115-115985	Sequence 115985, A	176	14.4	80.0	1968	17	US-10-437-963-18069	Sequence 18069, A
C 104	14.8	82.2	2701	17	US-10-437-963-36790	Sequence 36790, A	C 177	14.4	80.0	2157	18	US-10-425-115-127123	Sequence 127123, A
C 105	14.8	82.2	2712	16	US-10-282-122A-8910	Sequence 8910, Ap	178	14.4	80.0	2369	15	US-10-104-047-927	Sequence 927, Appl
C 106	14.8	82.2	2809	18	US-10-425-115-146651	Sequence 146651, A	179	14.4	80.0	2390	14	US-10-149-819-23	Sequence 819, App
C 107	14.8	82.2	2821	15	US-10-278-536-551	Sequence 53655, A	C 180	14.4	80.0	2400	16	US-10-424-599-23999	Sequence 23999, A
C 108	14.8	82.2	2821	15	US-10-225-068A-79	Sequence 51, Appl	181	14.4	80.0	2465	16	US-10-425-115-5589	Sequence 5889, App
C 109	14.8	82.2	2821	16	US-10-374-780A-77	Sequence 29, Appl	182	14.4	80.0	2670	17	US-10-437-963-14571	Sequence 12571, A
C 110	14.8	82.2	2821	16	US-10-412-699B-283	Sequence 283, App	183	14.4	80.0	2961	16	US-10-425-114-34064	Sequence 34064, A
C 111	14.8	82.2	2821	16	US-10-412-699B-283	Sequence 1775, Ap	184	14.4	80.0	3125	16	US-10-264-04A-868	Sequence 868, App
C 112	14.8	82.2	2831	16	US-10-425-115-114245	Sequence 114245, A	185	14.4	80.0	3296	9	US-09-909-0280-310	Sequence 310, App
C 113	14.8	82.2	3554	15	US-10-369-493-37050	Sequence 37050, A	186	14.4	80.0	3296	9	US-09-905-291A-310	Sequence 310, App
C 114	14.8	82.2	4854	16	US-10-282-122A-3314	Sequence 33184, A	187	14.4	80.0	3296	9	US-09-902-853-310	Sequence 310, App
C 115	14.8	82.2	6245	17	US-10-437-963-45500	Sequence 45500, A	188	14.4	80.0	3296	9	US-09-907-824-310	Sequence 310, App
C 116	14.8	82.2	40000	17	US-10-741-601-5780	Sequence 5780, App	189	14.4	80.0	3296	9	US-09-907-824-310	Sequence 310, App
C 117	14.8	82.2	60057	13	US-10-087-192-700	Sequence 700, App	190	14.4	80.0	3296	9	US-09-907-841-310	Sequence 310, App
C 118	14.8	82.2	84428	16	US-10-029-148B-1	Sequence 1, Appl	191	14.4	80.0	3296	9	US-09-904-011-310	Sequence 310, App
C 119	14.8	82.2	248436	13	US-10-087-192-2014	Sequence 2014, Ap	192	14.4	80.0	3296	10	US-09-903-640-310	Sequence 310, App
C 120	14.8	82.2	352938	17	US-10-323-656-79	Sequence 79, Appl	193	14.4	80.0	3296	10	US-09-908-090-310	Sequence 310, App
C 121	14.8	82.2	430442	18	US-10-427-437A-126	Sequence 128, App	194	14.4	80.0	3296	10	US-09-905-742-310	Sequence 310, App
C 122	14.4	80.0	60	15	US-10-272-437A-26	Sequence 26, Appl	195	14.4	80.0	3296	10	US-09-906-823-310	Sequence 310, App
C 123	14.4	80.0	174	15	US-10-029-386-19990	Sequence 19990, A	196	14.4	80.0	3296	10	US-09-907-613-310	Sequence 310, App
C 124	14.4	80.0	186	17	US-10-702-075-560	Sequence 560, App	197	14.4	80.0	3296	10	US-09-907-94-310	Sequence 310, App
C 125	14.4	80.0	192	17	US-10-437-963-4054	Sequence 4054, Ap	198	14.4	80.0	3296	10	US-09-904-853-310	Sequence 310, App
C 126	14.4	80.0	250	17	US-10-437-963-100379	Sequence 100379, A	199	14.4	80.0	3296	10	US-09-909-220-310	Sequence 310, App
C 127	14.4	80.0	257	18	US-10-425-115-44046	Sequence 44046, A	200	14.4	80.0	3296	10	US-09-904-824-310	Sequence 310, App
C 128	14.4	80.0	359	16	US-10-424-599-15793	Sequence 15793, A	201	14.4	80.0	3296	10	US-09-904-786-310	Sequence 310, App
C 129	14.4	80.0	424	15	US-10-425-115-16248	Sequence 16248, A	202	14.4	80.0	3296	10	US-09-906-54-310	Sequence 310, App
C 130	14.4	80.0	429	15	US-10-210-314-5	Sequence 5, Appl	203	14.4	80.0	3296	10	US-09-906-700-310	Sequence 310, App
C 131	14.4	80.0	458	18	US-10-425-115-40702	Sequence 40702, A	204	14.4	80.0	3296	10	US-09-903-788-310	Sequence 310, App
C 132	14.4	80.0	478	10	US-09-918-995-2414	Sequence 2414, Ap	205	14.4	80.0	3296	10	US-09-902-903-310	Sequence 310, App
C 133	14.4	80.0	493	10	US-09-918-995-31960	Sequence 31960, A	206	14.4	80.0	3296	10	US-09-903-749A-110	Sequence 310, App
C 134	14.4	80.0	502	17	US-10-021-323-1270	Sequence 1270, Ap	207	14.4	80.0	3296	10	US-09-904-119-310	Sequence 310, App
C 135	14.4	80.0	530	15	US-10-029-386-6259	Sequence 6259, Ap	208	14.4	80.0	3296	10	US-09-904-955-310	Sequence 310, App
C 136	14.4	80.0	532	17	US-10-767-701-30765	Sequence 30765, A	209	14.4	80.0	3296	10	US-09-902-735-310	Sequence 310, App
C 137	14.4	80.0	534	13	US-10-027-632-49968	Sequence 49968, A	210	14.4	80.0	3296	10	US-09-907-794-310	Sequence 310, App
C 138	14.4	80.0	553	15	US-10-027-632-49968	Sequence 49968, A	211	14.4	80.0	3296	10	US-09-903-943-310	Sequence 310, App
C 139	14.4	80.0	553	10	US-09-918-995-11478	Sequence 11478, A	212	14.4	80.0	3296	10	US-09-904-463-310	Sequence 310, App
C 140	14.4	80.0	569	13	US-10-027-632-81739	Sequence 81739, A	213	14.4	80.0	3296	10	US-09-907-922-310	Sequence 310, App
C 141	14.4	80.0	569	15	US-10-027-632-81739	Sequence 81739, A	214	14.4	80.0	3296	10	US-09-902-699-310	Sequence 310, App
C 142	14.4	80.0	591	17	US-10-479-334-2	Sequence 2, Appl	215	14.4	80.0	3296	10	US-09-903-520-310	Sequence 310, App
C 143	14.4	80.0	591	17	US-10-479-334-11	Sequence 11, Appl	216	14.4	80.0	3296	10	US-09-905-055-310	Sequence 310, App
C 144	14.4	80.0	596	16	US-10-424-599-99594	Sequence 39594, A	217	14.4	80.0	3296	10	US-09-909-064-310	Sequence 310, App
C 145	14.4	80.0	605	18	US-10-425-115-143653	Sequence 143653, A	218	14.4	80.0	3296	10	US-09-904-553-310	Sequence 310, App
C 146	14.4	80.0	650	16	US-10-424-599-9541	Sequence 9541, Ap	219	14.4	80.0	3296	10	US-09-905-381-310	Sequence 310, App
C 147	14.4	80.0	684	9	US-09-938-842A-963	Sequence 963, App	220	14.4	80.0	3296	10	US-09-904-485-310	Sequence 310, App
C 148	14.4	80.0	684	11	US-09-938-842A-963	Sequence 963, App	221	14.4	80.0	3296	10	US-09-905-381-310	Sequence 310, App
C 149	14.4	80.0	718	13	US-10-027-632-144355	Sequence 144355, A	222	14.4	80.0	3296	10	US-09-905-088-310	Sequence 310, App
C 150	14.4	80.0	718	13	US-10-027-632-144355	Sequence 144355, A	223	14.4	80.0	3296	10	US-09-907-577-310	Sequence 310, App
C 151	14.4	80.0	732	16	US-10-424-599-88952	Sequence 38952, A	224	14.4	80.0	3296	10	US-09-905-075-310	Sequence 310, App
C 152	14.4	80.0	748	17	US-10-767-701-7325	Sequence 7325, Ap	225	14.4	80.0	3296	10	US-09-902-759-310	Sequence 310, App
C 153	14.4	80.0	765	17	US-10-437-963-83366	Sequence 83366, A	226	14.4	80.0	3296	10	US-09-902-633-310	Sequence 310, App
C 154	14.4	80.0	802	16	US-10-424-599-5626	Sequence 5626, Ap	227	14.4	80.0	3296	10	US-09-902-713-310	Sequence 310, App
C 155	14.4	80.0	820	15	US-10-369-493-38770	Sequence 38770, A	228	14.4	80.0	3296	10	US-09-907-979-310	Sequence 310, App
C 156	14.4	80.0	843	17	US-10-437-963-18416	Sequence 18416, A	229	14.4	80.0	3296	10	US-09-902-615-310	Sequence 310, App
C 157	14.4	80.0	1116	15	US-10-369-493-32264	Sequence 32264, A	230	14.4	80.0	3296	10	US-09-903-925-310	Sequence 310, App
C 158	14.4	80.0	1159	11	US-09-985-153-45	Sequence 45, Appl	231	14.4	80.0	3296	10	US-09-906-760A-310	Sequence 310, App

232	14.4	80.0	3296	10	US-09-903-882-310	Sequence 310, App	305	14.4	80.0	3296	14	US-10-147-827-369	Sequence 369, App
233	14.4	80.0	3296	10	US-09-907-652-310	Sequence 310, App	306	14.4	80.0	3296	14	US-10-121-041-369	Sequence 369, App
234	14.4	80.0	3296	10	US-09-902-572A-310	Sequence 310, App	307	14.4	80.0	3296	14	US-10-111-043-369	Sequence 369, App
235	14.4	80.0	3296	10	US-09-902-979-310	Sequence 310, App	308	14.4	80.0	3296	14	US-10-121-015-369	Sequence 369, App
236	14.4	80.0	3296	10	US-09-905-125-310	Sequence 310, App	309	14.4	80.0	3296	14	US-10-123-002-369	Sequence 369, App
237	14.4	80.0	3296	10	US-09-906-815A-310	Sequence 310, App	310	14.4	80.0	3296	14	US-10-123-008-369	Sequence 369, App
238	14.4	80.0	3296	10	US-09-905-448-310	Sequence 310, App	311	14.4	80.0	3296	14	US-10-123-009-369	Sequence 369, App
239	14.4	80.0	3296	10	US-09-903-806-310	Sequence 310, App	312	14.4	80.0	3296	14	US-10-123-010-369	Sequence 369, App
240	14.4	80.0	3296	10	US-09-904-938-310	Sequence 310, App	313	14.4	80.0	3296	14	US-10-123-011-369	Sequence 369, App
241	14.4	80.0	3296	10	US-09-904-838-310	Sequence 310, App	314	14.4	80.0	3296	14	US-10-124-813-369	Sequence 369, App
242	14.4	80.0	3296	10	US-09-906-777-310	Sequence 310, App	315	14.4	80.0	3296	14	US-10-124-917-369	Sequence 369, App
243	14.4	80.0	3296	10	US-09-903-603A-310	Sequence 310, App	316	14.4	80.0	3296	14	US-10-125-922-369	Sequence 369, App
244	14.4	80.0	3296	10	US-09-904-532-310	Sequence 310, App	317	14.4	80.0	3296	14	US-10-125-924-369	Sequence 369, App
245	14.4	80.0	3296	10	US-09-904-766-310	Sequence 310, App	318	14.4	80.0	3296	14	US-10-140-860-369	Sequence 369, App
246	14.4	80.0	3296	10	US-09-904-928A-310	Sequence 310, App	319	14.4	80.0	3296	14	US-10-142-811-369	Sequence 369, App
247	14.4	80.0	3296	10	US-09-904-877A-310	Sequence 310, App	320	14.4	80.0	3296	14	US-10-147-819-369	Sequence 369, App
248	14.4	80.0	3296	10	US-09-903-562-310	Sequence 310, App	321	14.4	80.0	3296	14	US-10-137-782-369	Sequence 369, App
249	14.4	80.0	3296	10	US-09-906-618-310	Sequence 310, App	322	14.4	80.0	3296	14	US-10-152-959-369	Sequence 369, App
250	14.4	80.0	3296	10	US-09-907-728-310	Sequence 310, App	323	14.4	80.0	3296	14	US-10-135-926A-369	Sequence 369, App
251	14.4	80.0	3296	10	US-09-904-805-310	Sequence 310, App	324	14.4	80.0	3296	14	US-10-135-930A-369	Sequence 369, App
252	14.4	80.0	3296	10	US-09-904-938A-310	Sequence 310, App	325	14.4	80.0	3296	14	US-10-137-837A-369	Sequence 369, App
253	14.4	80.0	3296	10	US-09-906-722A-310	Sequence 310, App	326	14.4	80.0	3296	14	US-10-127-838A-369	Sequence 369, App
254	14.4	80.0	3296	11	US-09-908-576-310	Sequence 310, App	327	14.4	80.0	3296	14	US-10-117-882A-369	Sequence 369, App
255	14.4	80.0	3296	11	US-10-028-072-369	Sequence 369, App	328	14.4	80.0	3296	14	US-10-117-883A-369	Sequence 369, App
256	14.4	80.0	3296	14	US-10-140-808-369	Sequence 369, App	329	14.4	80.0	3296	14	US-10-127-843A-369	Sequence 369, App
257	14.4	80.0	3296	14	US-10-121-049-369	Sequence 369, App	330	14.4	80.0	3296	14	US-10-117-884A-369	Sequence 369, App
258	14.4	80.0	3296	14	US-10-123-904-369	Sequence 369, App	331	14.4	80.0	3296	14	US-10-137-846A-369	Sequence 369, App
259	14.4	80.0	3296	14	US-10-140-470-369	Sequence 369, App	332	14.4	80.0	3296	14	US-10-137-848A-369	Sequence 369, App
260	14.4	80.0	3296	14	US-10-175-746-369	Sequence 369, App	333	14.4	80.0	3296	14	US-10-137-849A-369	Sequence 369, App
261	14.4	80.0	3296	14	US-10-176-918-369	Sequence 369, App	334	14.4	80.0	3296	14	US-10-127-881A-369	Sequence 369, App
262	14.4	80.0	3296	14	US-10-176-921-369	Sequence 369, App	335	14.4	80.0	3296	14	US-10-127-882A-369	Sequence 369, App
263	14.4	80.0	3296	14	US-10-137-865-369	Sequence 369, App	336	14.4	80.0	3296	14	US-10-127-883A-369	Sequence 369, App
264	14.4	80.0	3296	14	US-10-140-474-369	Sequence 369, App	337	14.4	80.0	3296	14	US-10-128-666A-369	Sequence 369, App
265	14.4	80.0	3296	14	US-10-142-431-369	Sequence 369, App	338	14.4	80.0	3296	14	US-10-138-650A-369	Sequence 369, App
266	14.4	80.0	3296	14	US-10-143-114-369	Sequence 369, App	339	14.4	80.0	3296	14	US-10-138-651A-369	Sequence 369, App
267	14.4	80.0	3296	14	US-10-140-002-369	Sequence 369, App	340	14.4	80.0	3296	14	US-10-131-819A-369	Sequence 369, App
268	14.4	80.0	3296	14	US-10-142-419-369	Sequence 369, App	341	14.4	80.0	3296	14	US-10-131-820A-369	Sequence 369, App
269	14.4	80.0	3296	14	US-10-123-262-369	Sequence 369, App	342	14.4	80.0	3296	14	US-10-131-826A-369	Sequence 369, App
270	14.4	80.0	3296	14	US-10-142-423-369	Sequence 369, App	343	14.4	80.0	3296	14	US-10-146-729-369	Sequence 369, App
271	14.4	80.0	3296	14	US-10-121-050-369	Sequence 369, App	344	14.4	80.0	3296	14	US-10-146-791-369	Sequence 369, App
272	14.4	80.0	3296	14	US-10-141-755-369	Sequence 369, App	345	14.4	80.0	3296	14	US-10-147-484-369	Sequence 369, App
273	14.4	80.0	3296	14	US-10-143-032-369	Sequence 369, App	346	14.4	80.0	3296	14	US-10-147-508-369	Sequence 369, App
274	14.4	80.0	3296	14	US-10-123-108-369	Sequence 369, App	347	14.4	80.0	3296	14	US-10-147-512-369	Sequence 369, App
275	14.4	80.0	3296	14	US-10-123-263-369	Sequence 369, App	348	14.4	80.0	3296	14	US-10-175-735-369	Sequence 369, App
276	14.4	80.0	3296	14	US-10-123-261-369	Sequence 369, App	349	14.4	80.0	3296	14	US-10-121-040-369	Sequence 369, App
277	14.4	80.0	3296	14	US-10-140-921-369	Sequence 369, App	350	14.4	80.0	3296	14	US-10-121-056-369	Sequence 369, App
278	14.4	80.0	3296	14	US-10-140-928-369	Sequence 369, App	351	14.4	80.0	3296	14	US-10-121-061-369	Sequence 369, App
279	14.4	80.0	3296	14	US-10-121-045-369	Sequence 369, App	352	14.4	80.0	3296	14	US-10-123-235-369	Sequence 369, App
280	14.4	80.0	3296	14	US-10-123-232-369	Sequence 369, App	353	14.4	80.0	3296	14	US-10-124-811-369	Sequence 369, App
281	14.4	80.0	3296	14	US-10-123-903-369	Sequence 369, App	354	14.4	80.0	3296	14	US-10-137-868-369	Sequence 369, App
282	14.4	80.0	3296	14	US-10-124-811-369	Sequence 369, App	355	14.4	80.0	3296	14	US-10-147-492-369	Sequence 369, App
283	14.4	80.0	3296	14	US-10-124-821-369	Sequence 369, App	356	14.4	80.0	3296	14	US-10-138-782-369	Sequence 369, App
284	14.4	80.0	3296	14	US-10-140-925-369	Sequence 369, App	357	14.4	80.0	3296	14	US-10-123-905-369	Sequence 369, App
285	14.4	80.0	3296	14	US-10-160-498-369	Sequence 369, App	358	14.4	80.0	3296	14	US-10-127-836A-369	Sequence 369, App
286	14.4	80.0	3296	14	US-10-124-824-369	Sequence 369, App	359	14.4	80.0	3296	14	US-10-123-907-369	Sequence 369, App
287	14.4	80.0	3296	14	US-10-127-825A-369	Sequence 369, App	360	14.4	80.0	3296	14	US-10-124-815-369	Sequence 369, App
288	14.4	80.0	3296	14	US-10-127-822A-369	Sequence 369, App	361	14.4	80.0	3296	14	US-10-125-923A-369	Sequence 369, App
289	14.4	80.0	3296	14	US-10-127-832A-369	Sequence 369, App	362	14.4	80.0	3296	14	US-10-127-821A-369	Sequence 369, App
290	14.4	80.0	3296	14	US-10-127-839A-369	Sequence 369, App	363	14.4	80.0	3296	14	US-10-127-822A-369	Sequence 369, App
291	14.4	80.0	3296	14	US-10-127-901A-369	Sequence 369, App	364	14.4	80.0	3296	14	US-10-127-824A-369	Sequence 369, App
292	14.4	80.0	3296	14	US-10-128-693A-369	Sequence 369, App	365	14.4	80.0	3296	14	US-10-127-826A-369	Sequence 369, App
293	14.4	80.0	3296	14	US-10-131-818A-369	Sequence 369, App	366	14.4	80.0	3296	14	US-10-127-827A-369	Sequence 369, App
294	14.4	80.0	3296	14	US-10-131-815A-369	Sequence 369, App	367	14.4	80.0	3296	14	US-10-127-828A-369	Sequence 369, App
295	14.4	80.0	3296	14	US-10-131-823A-369	Sequence 369, App	368	14.4	80.0	3296	14	US-10-127-830A-369	Sequence 369, App
296	14.4	80.0	3296	14	US-10-131-824A-369	Sequence 369, App	369	14.4	80.0	3296	14	US-10-127-832A-369	Sequence 369, App
297	14.4	80.0	3296	14	US-10-131-824A-369	Sequence 369, App	370	14.4	80.0	3296	14	US-10-127-833A-369	Sequence 369, App
298	14.4	80.0	3296	14	US-10-131-830A-369	Sequence 369, App	371	14.4	80.0	3296	14	US-10-127-834A-369	Sequence 369, App
299	14.4	80.0	3296	14	US-10-131-837A-369	Sequence 369, App	372	14.4	80.0	3296	14	US-10-127-836A-369	Sequence 369, App
300	14.4	80.0	3296	14	US-10-137-872A-369	Sequence 369, App	373	14.4	80.0	3296	14	US-10-127-837A-369	Sequence 369, App
301	14.4	80.0	3296	14	US-10-147-500-369	Sequence 369, App	374	14.4	80.0	3296	14	US-10-127-841A-369	Sequence 369, App
302	14.4	80.0	3296	14	US-10-147-502-369	Sequence 369, App	375	14.4	80.0	3296	14	US-10-127-844A-369	Sequence 369, App
303	14.4	80.0	3296	14	US-10-147-515-369	Sequence 369, App	376	14.4	80.0	3296	14	US-10-128-687A-369	Sequence 369, App
304	14.4	80.0	3296	14	US-10-147-517-369	Sequence 369, App	377	14.4	80.0	3296	14	US-10-128-688A-369	Sequence 369, App
												US-10-128-699A-369	Sequence 369, App

378	14.4	80.0	3296	14	US-10-129-654A-369	Sequence 369, App	421	14.4	80.0	3296	15	US-10-141-700-369	Sequence 369, App
379	14.4	80.0	3296	14	US-10-131-825A-369	Sequence 369, App	422	14.4	80.0	3296	15	US-10-141-705-369	Sequence 369, App
380	14.4	80.0	3296	14	US-10-230-417-369	Sequence 369, App	423	14.4	80.0	3296	15	US-10-141-753-369	Sequence 369, App
381	14.4	80.0	3296	14	US-10-131-815A-369	Sequence 369, App	424	14.4	80.0	3296	15	US-10-141-758-369	Sequence 369, App
382	14.4	80.0	3296	14	US-10-131-817A-369	Sequence 369, App	425	14.4	80.0	3296	15	US-10-142-420-369	Sequence 369, App
383	14.4	80.0	3296	14	US-10-131-821A-369	Sequence 369, App	426	14.4	80.0	3296	15	US-10-142-422-369	Sequence 369, App
384	14.4	80.0	3296	14	US-10-131-822A-369	Sequence 369, App	427	14.4	80.0	3296	15	US-10-142-423-369	Sequence 369, App
385	14.4	80.0	3296	14	US-10-131-828A-369	Sequence 369, App	428	14.4	80.0	3296	15	US-10-142-760-369	Sequence 369, App
386	14.4	80.0	3296	14	US-10-131-835A-369	Sequence 369, App	429	14.4	80.0	3296	15	US-10-142-821-369	Sequence 369, App
387	14.4	80.0	3296	14	US-10-137-864A-369	Sequence 369, App	430	14.4	80.0	3296	15	US-10-152-531-369	Sequence 369, App
388	14.4	80.0	3296	14	US-10-137-869A-369	Sequence 369, App	431	14.4	80.0	3296	15	US-10-152-531-369	Sequence 369, App
389	14.4	80.0	3296	14	US-10-147-523-369	Sequence 369, App	432	14.4	80.0	3296	15	US-10-142-424-369	Sequence 369, App
390	14.4	80.0	3296	14	US-10-158-785-369	Sequence 369, App	433	14.4	80.0	3296	15	US-10-142-424-369	Sequence 369, App
391	14.4	80.0	3296	14	US-10-121-051-369	Sequence 369, App	434	14.4	80.0	3296	15	US-10-142-761-369	Sequence 369, App
392	14.4	80.0	3296	14	US-10-121-042-369	Sequence 369, App	435	14.4	80.0	3296	15	US-10-142-763-369	Sequence 369, App
393	14.4	80.0	3296	15	US-10-123-912-369	Sequence 369, App	436	14.4	80.0	3296	15	US-10-144-957-369	Sequence 369, App
394	14.4	80.0	3296	15	US-10-192-007-369	Sequence 369, App	437	14.4	80.0	3296	15	US-10-144-957-369	Sequence 369, App
395	14.4	80.0	3296	15	US-10-194-359-369	Sequence 369, App	438	14.4	80.0	3296	15	US-10-142-887-369	Sequence 369, App
396	14.4	80.0	3296	15	US-10-194-359-369	Sequence 369, App	439	14.4	80.0	3296	15	US-10-142-888-369	Sequence 369, App
397	14.4	80.0	3296	15	US-10-127-847A-369	Sequence 369, App	440	14.4	80.0	3296	15	US-10-143-034-369	Sequence 369, App
398	14.4	80.0	3296	15	US-10-175-860-369	Sequence 369, App	441	14.4	80.0	3296	15	US-10-143-116-369	Sequence 369, App
399	14.4	80.0	3296	15	US-10-137-866-369	Sequence 369, App	442	14.4	80.0	3296	15	US-10-143-117-369	Sequence 369, App
399	14.4	80.0	3296	15	US-10-146-726-369	Sequence 369, App	442	14.4	80.0	3296	15	US-10-144-957-369	Sequence 369, App
400	14.4	80.0	3296	15	US-10-146-777-369	Sequence 369, App	442	14.4	80.0	3296	15	US-10-144-957-369	Sequence 369, App
401	14.4	80.0	3296	15	US-10-146-788-369	Sequence 369, App	443	14.4	80.0	3296	15	US-10-144-957-369	Sequence 369, App
402	14.4	80.0	3296	15	US-10-146-788-369	Sequence 369, App	443	14.4	80.0	3296	15	US-10-145-015-369	Sequence 369, App
402	14.4	80.0	3296	15	US-10-152-380-369	Sequence 369, App	445	14.4	80.0	3296	15	US-10-145-015-369	Sequence 369, App
403	14.4	80.0	3296	15	US-10-153-934-369	Sequence 369, App	445	14.4	80.0	3296	15	US-10-145-019-369	Sequence 369, App
403	14.4	80.0	3296	15	US-10-153-934-369	Sequence 369, App	445	14.4	80.0	3296	15	US-10-145-019-369	Sequence 369, App
404	14.4	80.0	3296	15	US-10-140-807-369	Sequence 369, App	447	14.4	80.0	3296	15	US-10-145-021-369	Sequence 369, App
404	14.4	80.0	3296	15	US-10-140-807-369	Sequence 369, App	447	14.4	80.0	3296	15	US-10-145-021-369	Sequence 369, App
405	14.4	80.0	3296	15	US-10-140-924-369	Sequence 369, App	447	14.4	80.0	3296	15	US-10-145-021-369	Sequence 369, App
406	14.4	80.0	3296	15	US-10-140-924-369	Sequence 369, App	447	14.4	80.0	3296	15	US-10-145-021-369	Sequence 369, App
407	14.4	80.0	3296	15	US-10-140-926-369	Sequence 369, App	447	14.4	80.0	3296	15	US-10-145-021-369	Sequence 369, App
407	14.4	80.0	3296	15	US-10-141-698-369	Sequence 369, App	447	14.4	80.0	3296	15	US-10-145-021-369	Sequence 369, App
408	14.4	80.0	3296	15	US-10-141-702-369	Sequence 369, App	448	14.4	80.0	3296	15	US-10-145-021-369	Sequence 369, App
409	14.4	80.0	3296	15	US-10-141-702-369	Sequence 369, App	448	14.4	80.0	3296	15	US-10-145-021-369	Sequence 369, App
410	14.4	80.0	3296	15	US-10-141-704-369	Sequence 369, App	448	14.4	80.0	3296	15	US-10-145-021-369	Sequence 369, App
410	14.4	80.0	3296	15	US-10-142-421-369	Sequence 369, App	448	14.4	80.0	3296	15	US-10-145-021-369	Sequence 369, App
411	14.4	80.0	3296	15	US-10-142-432-369	Sequence 369, App	448	14.4	80.0	3296	15	US-10-145-021-369	Sequence 369, App
412	14.4	80.0	3296	15	US-10-142-432-369	Sequence 369, App	448	14.4	80.0	3296	15	US-10-145-021-369	Sequence 369, App
412	14.4	80.0	3296	15	US-10-142-432-369	Sequence 369, App	448	14.4	80.0	3296	15	US-10-145-021-369	Sequence 369, App
413	14.4	80.0	3296	15	US-10-143-013-369	Sequence 369, App	448	14.4	80.0	3296	15	US-10-145-021-369	Sequence 369, App
413	14.4	80.0	3296	15	US-10-143-013-369	Sequence 369, App	448	14.4	80.0	3296	15	US-10-145-021-369	Sequence 369, App
414	14.4	80.0	3296	15	US-10-144-994-369	Sequence 369, App	448	14.4	80.0	3296	15	US-10-145-021-369	Sequence 369, App
414	14.4	80.0	3296	15	US-10-144-994-369	Sequence 369, App	448	14.4	80.0	3296	15	US-10-145-021-369	Sequence 369, App
415	14.4	80.0	3296	15	US-10-145-628-369	Sequence 369, App	448	14.4	80.0	3296	15	US-10-145-021-369	Sequence 369, App
415	14.4	80.0	3296	15	US-10-145-628-369	Sequence 369, App	448	14.4	80.0	3296	15	US-10-145-021-369	Sequence 369, App
416	14.4	80.0	3296	15	US-10-145-746-369	Sequence 369, App	448	14.4	80.0	3296	15	US-10-145-021-369	Sequence 369, App
417	14.4	80.0	3296	15	US-10-145-746-369	Sequence 369, App	448	14.4	80.0	3296	15	US-10-145-021-369	Sequence 369, App
418	14.4	80.0	3296	15	US-10-145-823-369	Sequence 369, App	448	14.4	80.0	3296	15	US-10-145-021-369	Sequence 369, App
419	14.4	80.0	3296	15	US-10-145-823-369	Sequence 369, App	448	14.4	80.0	3296	15	US-10-145-021-369	Sequence 369, App
419	14.4	80.0	3296	15	US-10-145-826-369	Sequence 369, App	448	14.4	80.0	3296	15	US-10-145-021-369	Sequence 369, App
420	14.4	80.0	3296	15	US-10-145-870-369	Sequence 369, App	448	14.4	80.0	3296	15	US-10-145-021-369	Sequence 369, App
421	14.4	80.0	3296	15	US-10-145-870-369	Sequence 369, App	448	14.4	80.0	3296	15	US-10-145-021-369	Sequence 369, App
422	14.4	80.0	3296	15	US-10-145-876-369	Sequence 369, App	448	14.4	80.0	3296	15	US-10-145-021-369	Sequence 369, App
422	14.4	80.0	3296	15	US-10-145-876-369	Sequence 369, App	448	14.4	80.0	3296	15	US-10-145-021-369	Sequence 369, App
423	14.4	80.0	3296	15	US-10-145-959-369	Sequence 369, App	448	14.4	80.0	3296	15	US-10-145-021-369	Sequence 369, App
423	14.4	80.0	3296	15	US-10-145-959-369	Sequence 369, App	448	14.4	80.0	3296	15	US-10-145-021-369	Sequence 369, App
424	14.4	80.0	3296	15	US-10-146-724-369	Sequence 369, App	448	14.4	80.0	3296	15	US-10-145-021-369	Sequence 369, App
424	14.4	80.0	3296	15	US-10-146-724-369	Sequence 369, App	448	14.4	80.0	3296	15	US-10-145-021-369	Sequence 369, App
425	14.4	80.0	3296	15	US-10-146-725-369	Sequence 369, App	448	14.4	80.0	3296	15	US-10-145-021-369	Sequence 369, App
425	14.4	80.0	3296	15	US-10-146-725-369	Sequence 369, App	448	14.4	80.0	3296	15	US-10-145-021-369	Sequence 369, App
426	14.4	80.0	3296	15	US-10-146-795-369	Sequence 369, App	448	14.4	80.0	3296	15	US-10-145-021-369	Sequence 369, App
427	14.4	80.0	3296	15	US-10-147-495-369	Sequence 369, App	448	14.4	80.0	3296	15	US-10-145-021-369	Sequence 369, App
427	14.4	80.0	3296	15	US-10-147-495-369	Sequence 369, App	448	14.4	80.0	3296	15	US-10-145-021-369	Sequence 369, App
428	14.4	80.0	3296	15	US-10-147-501-369	Sequence 369, App	448	14.4	80.0	3296	15	US-10-145-021-369	Sequence 369, App
429	14.4	80.0	3296	15	US-10-147-501-369	Sequence 369, App	448	14.4	80.0	3296	15	US-10-145-021-369	Sequence 369, App
429	14.4	80.0	3296	15	US-10-147-504-369	Sequence 369, App	448	14.4	80.0	3296	15	US-10-145-021-369	Sequence 369, App
430	14.4	80.0	3296	15	US-10-147-504-369	Sequence 369, App	448	14.4	80.0	3296	15	US-10-145-021-369	Sequence 369, App
431	14.4	80.0	3296	15	US-10-147-509-369	Sequence 369, App	448	14.4	80.0	3296	15	US-10-145-021-369	Sequence 369, App
431	14.4	80.0	3296	15	US-10-147-509-369	Sequence 369, App	448	14.4	80.0	3296	15	US-10-145-021-369	Sequence 369, App
432	14.4	80.0	3296	15	US-10-147-510-369	Sequence 369, App	448	14.4	80.0	3296	15	US-10-145-021-369	Sequence 369, App
432	14.4	80.0	3296	15	US-10-147-510-369	Sequence 369, App	448	14.4	80.0	3296	15	US-10-145-021-369	Sequence 369, App
433	14.4	80.0	3296	15	US-10-147-511-369	Sequence 369, App	448	14.4	80.0	3296	15	US-10-145-021-369	Sequence 369, App
433	14.4	80.0	3296	15	US-10-147-529-369	Sequence 369, App	448	14.4	80.0	3296	15	US-10-145-021-369	Sequence 369, App
434	14.4	80.0	3296	15	US-10-152-337-369	Sequence 369, App	448	14.4	80.0	3296	15	US-10-145-021-369	Sequence 369, App
435	14.4	80.0	3296	15	US-10-153-586-369	Sequence 369, App	448	14.4	80.0	3296	15	US-10-145-021-369	Sequence 369, App
436	14.4	80.0	3296	15	US-10-158-786-369	Sequence 369, App	448	14.4	80.0	3296	15	US-10-145-021-369	Sequence 369, App
437	14.4	80.0	3296	15	US-10-137-870-369	Sequence 369, App	448	14.4	80.0	3296	15	US-10-145-021-369	Sequence 369, App
438	14.4	80.0	3296	15	US-10-								

524	14.4	80.0	3296	15	US-10-147-520-369	Sequence 369, App	597	14.4	80.0	3296	15	US-10-153-552-369	Sequence 369, App
525	14.4	80.0	3296	15	US-10-157-781-369	Sequence 369, App	598	14.4	80.0	3296	15	US-10-153-840-369	Sequence 369, App
526	14.4	80.0	3296	15	US-10-176-989-369	Sequence 369, App	599	14.4	80.0	3296	15	US-10-153-841-369	Sequence 369, App
527	14.4	80.0	3296	15	US-10-147-491-369	Sequence 369, App	600	14.4	80.0	3296	15	US-10-156-842-369	Sequence 369, App
528	14.4	80.0	3296	15	US-10-152-378-369	Sequence 369, App	601	14.4	80.0	3296	15	US-10-156-844-369	Sequence 369, App
529	14.4	80.0	3296	15	US-10-152-382-369	Sequence 369, App	602	14.4	80.0	3296	15	US-10-156-845-369	Sequence 369, App
530	14.4	80.0	3296	15	US-10-152-383-369	Sequence 369, App	603	14.4	80.0	3296	15	US-10-156-846-369	Sequence 369, App
531	14.4	80.0	3296	15	US-10-152-384-369	Sequence 369, App	604	14.4	80.0	3296	15	US-10-121-048-369	Sequence 369, App
532	14.4	80.0	3296	15	US-10-152-387-369	Sequence 369, App	605	14.4	80.0	3296	15	US-10-121-052-369	Sequence 369, App
533	14.4	80.0	3296	15	US-10-152-389-369	Sequence 369, App	606	14.4	80.0	3296	15	US-10-121-053-369	Sequence 369, App
534	14.4	80.0	3296	15	US-10-152-390-369	Sequence 369, App	607	14.4	80.0	3296	15	US-10-121-054-369	Sequence 369, App
535	14.4	80.0	3296	15	US-10-152-392-369	Sequence 369, App	608	14.4	80.0	3296	15	US-10-123-412-369	Sequence 369, App
536	14.4	80.0	3296	15	US-10-152-403-369	Sequence 369, App	609	14.4	80.0	3296	15	US-10-123-413-369	Sequence 369, App
537	14.4	80.0	3296	15	US-10-153-756-369	Sequence 369, App	610	14.4	80.0	3296	15	US-10-123-291-369	Sequence 369, App
538	14.4	80.0	3296	15	US-10-157-784-369	Sequence 369, App	611	14.4	80.0	3296	15	US-10-123-322-369	Sequence 369, App
539	14.4	80.0	3296	15	US-10-157-797-369	Sequence 369, App	612	14.4	80.0	3296	15	US-10-123-771-369	Sequence 369, App
540	14.4	80.0	3296	15	US-10-158-491-369	Sequence 369, App	613	14.4	80.0	3296	15	US-10-123-911-369	Sequence 369, App
541	14.4	80.0	3296	15	US-10-239-976-310	Sequence 310, App	614	14.4	80.0	3296	15	US-10-124-820-369	Sequence 369, App
542	14.4	80.0	3296	15	US-10-142-762-369	Sequence 369, App	615	14.4	80.0	3296	15	US-10-124-823-369	Sequence 369, App
543	14.4	80.0	3296	15	US-10-142-764-369	Sequence 369, App	616	14.4	80.0	3296	15	US-10-125-931-369	Sequence 369, App
544	14.4	80.0	3296	15	US-10-142-766-369	Sequence 369, App	617	14.4	80.0	3296	15	US-10-127-852A-369	Sequence 369, App
545	14.4	80.0	3296	15	US-10-145-625-369	Sequence 369, App	618	14.4	80.0	3296	15	US-10-127-500A-369	Sequence 369, App
546	14.4	80.0	3296	15	US-10-145-627-369	Sequence 369, App	619	14.4	80.0	3296	15	US-10-128-685A-369	Sequence 369, App
547	14.4	80.0	3296	15	US-10-145-960-369	Sequence 369, App	620	14.4	80.0	3296	15	US-10-131-820A-369	Sequence 369, App
548	14.4	80.0	3296	15	US-10-145-962-369	Sequence 369, App	621	14.4	80.0	3296	15	US-10-142-686-369	Sequence 369, App
549	14.4	80.0	3296	15	US-10-146-789-369	Sequence 369, App	622	14.4	80.0	3296	15	US-10-142-686-369	Sequence 369, App
550	14.4	80.0	3296	15	US-10-147-483-369	Sequence 369, App	623	14.4	80.0	3296	15	US-10-146-728-369	Sequence 369, App
551	14.4	80.0	3296	15	US-10-147-486-369	Sequence 369, App	624	14.4	80.0	3296	15	US-10-146-786-369	Sequence 369, App
552	14.4	80.0	3296	15	US-10-147-505-369	Sequence 369, App	625	14.4	80.0	3296	15	US-10-147-499-369	Sequence 369, App
553	14.4	80.0	3296	15	US-10-147-516-369	Sequence 369, App	626	14.4	80.0	3296	15	US-10-147-798-369	Sequence 369, App
554	14.4	80.0	3296	15	US-10-153-398-369	Sequence 369, App	627	14.4	80.0	3296	15	US-10-147-798-369	Sequence 369, App
555	14.4	80.0	3296	15	US-10-139-980-369	Sequence 369, App	628	14.4	80.0	3296	15	US-10-147-798-369	Sequence 369, App
556	14.4	80.0	3296	15	US-10-299-937-310	Sequence 310, App	629	14.4	80.0	3296	15	US-10-137-667-369	Sequence 369, App
557	14.4	80.0	3296	15	US-10-145-750-369	Sequence 369, App	630	14.4	80.0	3296	15	US-10-140-473-369	Sequence 369, App
558	14.4	80.0	3296	15	US-10-152-373-369	Sequence 369, App	631	14.4	80.0	3296	15	US-10-140-810-369	Sequence 369, App
559	14.4	80.0	3296	15	US-10-121-044-369	Sequence 369, App	632	14.4	80.0	3296	15	US-10-140-863-369	Sequence 369, App
560	14.4	80.0	3296	15	US-10-121-055-369	Sequence 369, App	633	14.4	80.0	3296	15	US-10-141-699-369	Sequence 369, App
561	14.4	80.0	3296	15	US-10-121-057-369	Sequence 369, App	634	14.4	80.0	3296	15	US-10-141-703-369	Sequence 369, App
562	14.4	80.0	3296	15	US-10-121-058-369	Sequence 369, App	635	14.4	80.0	3296	15	US-10-141-706-369	Sequence 369, App
563	14.4	80.0	3296	15	US-10-121-059-369	Sequence 369, App	636	14.4	80.0	3296	15	US-10-141-757-369	Sequence 369, App
564	14.4	80.0	3296	15	US-10-121-060-369	Sequence 369, App	637	14.4	80.0	3296	15	US-10-141-762-369	Sequence 369, App
565	14.4	80.0	3296	15	US-10-123-109-369	Sequence 369, App	638	14.4	80.0	3296	15	US-10-142-428-369	Sequence 369, App
566	14.4	80.0	3296	15	US-10-123-134-369	Sequence 369, App	639	14.4	80.0	3296	15	US-10-142-429-369	Sequence 369, App
567	14.4	80.0	3296	15	US-10-123-157-369	Sequence 369, App	640	14.4	80.0	3296	15	US-10-142-884-369	Sequence 369, App
568	14.4	80.0	3296	15	US-10-123-190-369	Sequence 369, App	641	14.4	80.0	3296	15	US-10-143-027-369	Sequence 369, App
569	14.4	80.0	3296	15	US-10-124-814-369	Sequence 369, App	642	14.4	80.0	3296	15	US-10-143-115-369	Sequence 369, App
570	14.4	80.0	3296	15	US-10-124-816-369	Sequence 369, App	643	14.4	80.0	3296	15	US-10-144-356-369	Sequence 369, App
571	14.4	80.0	3296	15	US-10-124-820-369	Sequence 369, App	644	14.4	80.0	3296	15	US-10-144-358-369	Sequence 369, App
572	14.4	80.0	3296	15	US-10-125-704-369	Sequence 369, App	645	14.4	80.0	3296	15	US-10-145-632-369	Sequence 369, App
573	14.4	80.0	3296	15	US-10-125-927-369	Sequence 369, App	646	14.4	80.0	3296	15	US-10-145-749-369	Sequence 369, App
574	14.4	80.0	3296	15	US-10-125-974-369	Sequence 369, App	647	14.4	80.0	3296	15	US-10-145-753-369	Sequence 369, App
575	14.4	80.0	3296	15	US-10-145-879-369	Sequence 369, App	648	14.4	80.0	3296	15	US-10-145-871-369	Sequence 369, App
576	14.4	80.0	3296	15	US-10-145-884-369	Sequence 369, App	649	14.4	80.0	3296	15	US-10-145-878-369	Sequence 369, App
577	14.4	80.0	3296	15	US-10-147-497-369	Sequence 369, App	650	14.4	80.0	3296	15	US-10-146-794-369	Sequence 369, App
578	14.4	80.0	3296	15	US-10-152-374-369	Sequence 369, App	651	14.4	80.0	3296	15	US-10-147-489-369	Sequence 369, App
579	14.4	80.0	3296	15	US-10-152-375-369	Sequence 369, App	652	14.4	80.0	3296	15	US-10-147-489-369	Sequence 369, App
580	14.4	80.0	3296	15	US-10-153-331-369	Sequence 369, App	653	14.4	80.0	3296	15	US-10-147-535-369	Sequence 369, App
581	14.4	80.0	3296	15	US-10-153-336-369	Sequence 369, App	654	14.4	80.0	3296	15	US-10-147-537-369	Sequence 369, App
582	14.4	80.0	3296	15	US-10-153-331-369	Sequence 369, App	655	14.4	80.0	3296	15	US-10-147-376-369	Sequence 369, App
583	14.4	80.0	3296	15	US-10-153-339-369	Sequence 369, App	656	14.4	80.0	3296	15	US-10-152-381-369	Sequence 369, App
584	14.4	80.0	3296	15	US-10-153-339-369	Sequence 369, App	657	14.4	80.0	3296	15	US-10-152-400-369	Sequence 369, App
585	14.4	80.0	3296	15	US-10-157-785-369	Sequence 369, App	658	14.4	80.0	3296	15	US-10-153-585-369	Sequence 369, App
586	14.4	80.0	3296	15	US-10-157-786-369	Sequence 369, App	659	14.4	80.0	3296	15	US-10-157-780-369	Sequence 369, App
587	14.4	80.0	3296	15	US-10-157-794-369	Sequence 369, App	660	14.4	80.0	3296	15	US-10-157-800-369	Sequence 369, App
588	14.4	80.0	3296	15	US-10-160-500-369	Sequence 369, App	661	14.4	80.0	3296	15	US-10-157-801-369	Sequence 369, App
589	14.4	80.0	3296	15	US-10-121-046-369	Sequence 369, App	662	14.4	80.0	3296	15	US-10-157-802-369	Sequence 369, App
590	14.4	80.0	3296	15	US-10-123-156-369	Sequence 369, App	663	14.4	80.0	3296	15	US-10-158-784-369	Sequence 369, App
591	14.4	80.0	3296	15	US-10-123-214-369	Sequence 369, App	664	14.4	80.0	3296	15	US-10-158-789-369	Sequence 369, App
592	14.4	80.0	3296	15	US-10-123-805-369	Sequence 369, App	665	14.4	80.0	3296	15	US-10-132-011-369	Sequence 369, App
593	14.4	80.0	3296	15	US-10-124-821-369	Sequence 369, App	666	14.4	80.0	3296	15	US-10-139-963-369	Sequence 369, App
594	14.4	80.0	3296	15	US-10-152-385-369	Sequence 369, App	667	14.4	80.0	3296	15	US-10-140-020-369	Sequence 369, App
595	14.4	80.0	3296	15	US-10-152-393-369	Sequence 369, App	668	14.4	80.0	3296	15	US-10-140-023-369	Sequence 369, App
596	14.4	80.0	3296	15	US-10-153-366-369	Sequence 369, App	669	14.4	80.0	3296	15	US-10-140-809-369	Sequence 369, App

670	14.4	80.0	3296	15	US-10-140-865-369	Sequence 369, App	743	14	77.8	875	16	US-10-424-599-29925	Sequence 29925, A
671	14.4	80.0	3296	15	US-10-141-701-369	Sequence 369, App	744	14	77.8	892	17	US-10-437-963-92960	Sequence 92960, A
672	14.4	80.0	3296	15	US-10-141-754-369	Sequence 369, App	745	14	77.8	1022	16	US-10-424-599-4182	Sequence 4182, A
673	14.4	80.0	3296	15	US-10-141-760-369	Sequence 369, App	746	14	77.8	1041	17	US-10-437-963-71730	Sequence 71730, A
674	14.4	80.0	3296	15	US-10-142-425-369	Sequence 369, App	747	14	77.8	1209	15	US-10-032-583-6634	Sequence 6634, A
675	14.4	80.0	3296	15	US-10-142-430-369	Sequence 369, App	748	14	77.8	1515	15	US-10-295-400-141	Sequence 141, A
676	14.4	80.0	3296	15	US-10-143-113-369	Sequence 369, App	749	14	77.8	1515	16	US-10-412-699B-535	Sequence 535, A
677	14.4	80.0	3296	15	US-10-146-730-369	Sequence 369, App	750	14	77.8	1653	16	US-10-282-122A-18952	Sequence 18952, A
678	14.4	80.0	3296	15	US-10-146-792-369	Sequence 369, App	751	14	77.8	2207	17	US-10-437-963-30562	Sequence 30562, A
679	14.4	80.0	3296	15	US-10-158-791-369	Sequence 369, App	752	14	77.8	2217	17	US-10-437-963-992950	Sequence 92950, A
680	14.4	80.0	3296	15	US-10-158-843-369	Sequence 369, App	753	14	77.8	2628	11	US-09-809-663A-19	Sequence 39, App
681	14.4	80.0	3296	15	US-10-157-788-369	Sequence 369, App	754	14	77.8	2833	17	US-10-437-963-71163	Sequence 73163, A
682	14.4	80.0	3296	15	US-10-152-405-369	Sequence 369, App	755	14	77.8	3144	17	US-10-437-963-71167	Sequence 73167, A
683	14.4	80.0	3296	15	US-10-298-993-310	Sequence 310, App	756	14	77.8	4828	15	US-10-437-963-80480	Sequence 80480, A
684	14.4	80.0	3296	15	US-10-147-528-369	Sequence 369, App	757	14	77.8	6455	17	US-10-017-161-963	Sequence 963, App
685	14.4	80.0	3296	15	US-10-448-923-310	Sequence 310, App	758	14	77.8	6678	18	US-10-739-930-106	Sequence 106, App
686	14.4	80.0	3296	15	US-10-448-923-310	Sequence 310, App	759	14	77.8	27459	15	US-10-017-161-1015	Sequence 1015, App
687	14.4	80.0	3296	16	US-10-448-713-310	Sequence 310, App	760	14	77.8	53885	13	US-10-087-199-1144	Sequence 1744, App
688	14.4	80.0	3296	16	US-10-128-692A-169	Sequence 369, App	761	14	77.8	51288	17	US-10-322-281-112	Sequence 112, App
689	14.4	80.0	3296	16	US-10-140-927-369	Sequence 369, App	762	14	77.8	91278	17	US-10-322-281-341	Sequence 341, App
690	14.4	80.0	3296	16	US-10-140-927-369	Sequence 310, App	763	14	77.8	118931	13	US-10-087-195-1108	Sequence 1108, App
691	14.4	80.0	3296	16	US-10-425-447-310	Sequence 369, App	764	14	77.8	256190	17	US-10-322-281-330	Sequence 320, App
692	14.4	80.0	3296	16	US-10-145-127-369	Sequence 369, App	765	14	77.8	659158	9	US-09-771-208-20	Sequence 20, App
693	14.4	80.0	3296	16	US-10-160-503-369	Sequence 369, App	766	14	77.8	2140405	13	US-10-027-632-76212	Sequence 76212, A
694	14.4	80.0	3296	16	US-10-143-118-369	Sequence 369, App	767	14	77.8	2140405	15	US-10-027-632-76212	Sequence 76212, A
695	14.4	80.0	3296	16	US-10-144-993-369	Sequence 369, App	768	14	77.8	240917	13	US-10-027-632-174763	Sequence 174763, A
696	14.4	80.0	3296	16	US-10-158-787-369	Sequence 369, App	769	14	77.8	2940917	15	US-10-027-632-174763	Sequence 174763, A
697	14.4	80.0	3296	16	US-10-140-024-369	Sequence 369, App	770	13.8	76.7	66	6	US-08-781-986A-1643	Sequence 1643, App
698	14.4	80.0	3296	16	US-10-147-536-369	Sequence 369, App	771	13.8	76.7	66	6	US-10-329-624-1643	Sequence 1643, App
699	14.4	80.0	3296	17	US-10-152-372-369	Sequence 369, App	772	13.8	76.7	159	16	US-10-282-122A-1569	Sequence 1569, App
700	14.4	80.0	3296	17	US-10-215-371-310	Sequence 310, App	773	13.8	76.7	160	9	US-09-604-287A-44	Sequence 44, App
701	14.4	80.0	3296	17	US-10-771-187-310	Sequence 310, App	774	13.8	76.7	160	9	US-09-339-338-44	Sequence 44, App
702	14.4	80.0	3296	18	US-10-125-795-369	Sequence 369, App	775	13.8	76.7	160	10	US-09-551-621-44	Sequence 44, App
703	14.4	80.0	3296	18	US-10-145-819-369	Sequence 369, App	776	13.8	76.7	160	10	US-10-007-805-44	Sequence 44, App
704	14.4	80.0	3296	18	US-10-145-819-369	Sequence 369, App	777	13.8	76.7	160	14	US-10-076-622-44	Sequence 44, App
705	14.4	80.0	3296	18	US-10-145-825-369	Sequence 369, App	778	13.8	76.7	160	15	US-10-124-805-44	Sequence 44, App
706	14.4	80.0	3296	18	US-10-147-513-369	Sequence 369, App	779	13.8	76.7	160	15	US-10-441-893-44	Sequence 44, App
707	14.4	80.0	3296	18	US-10-147-518-369	Sequence 369, App	780	13.8	76.7	170	9	US-09-815-242-3038	Sequence 3028, App
708	14.4	80.0	3296	18	US-10-145-961-369	Sequence 369, App	781	13.8	76.7	170	6	US-10-425-115-15452	Sequence 15452, A
709	14.4	80.0	3359	17	US-10-437-963-78567	Sequence 78567, A	782	13.8	76.7	170	16	US-10-282-122A-5592	Sequence 5592, A
710	14.4	80.0	3359	17	US-10-437-963-12579	Sequence 12579, A	783	13.8	76.7	171	18	US-10-425-115-182683	Sequence 182683, A
711	14.4	80.0	3726	17	US-10-437-963-13215	Sequence 13215, A	784	13.8	76.7	205	17	US-10-437-963-10642	Sequence 10642, A
712	14.4	80.0	3947	18	US-10-425-115-152026	Sequence 152026, A	785	13.8	76.7	223	9	US-09-983-965-1371	Sequence 1371, App
713	14.4	80.0	3999	17	US-10-437-963-35352	Sequence 35352, A	786	13.8	76.7	247	9	US-09-834-975-537	Sequence 537, App
714	14.4	80.0	4033	18	US-10-425-115-11862	Sequence 33862, A	787	13.8	76.7	250	8	US-10-425-115-15452	Sequence 15452, A
715	14.4	80.0	4426	16	US-10-094-886-123	Sequence 123, App	788	13.8	76.7	251	16	US-10-242-538A-14373	Sequence 14373, A
716	14.4	80.0	4488	8	US-08-781-986A-228	Sequence 228, App	789	13.8	76.7	251	16	US-10-085-738A-14373	Sequence 14373, A
717	14.4	80.0	4488	16	US-10-323-624-228	Sequence 228, App	790	13.8	76.7	255	16	US-10-282-122A-5597	Sequence 5597, A
718	14.4	80.0	4501	16	US-09-764-886-125	Sequence 125, App	791	13.8	76.7	267	14	US-10-046-433-10	Sequence 433, A
719	14.4	80.0	12123	10	US-09-764-872-752	Sequence 752, App	792	13.8	76.7	267	14	US-10-046-433-51	Sequence 433, A
720	14.4	80.0	35183	17	US-10-323-281-445	Sequence 445, App	793	13.8	76.7	288	9	US-09-983-965-20047	Sequence 20047, A
721	14.4	80.0	89213	17	US-10-323-281-843	Sequence 843, App	794	13.8	76.7	305	15	US-10-029-368-20047	Sequence 20047, A
722	14.4	80.0	119586	14	US-10-270-336-3	Sequence 3, App	795	13.8	76.7	317	16	US-10-282-122A-1819	Sequence 2819, App
723	14.4	80.0	162450	15	US-10-071-179-1	Sequence 1, App	796	13.8	76.7	322	16	US-10-282-122A-1819	Sequence 2819, App
724	14.4	80.0	162450	15	US-10-126-704-1	Sequence 1, App	797	13.8	76.7	327	18	US-10-425-115-4451	Sequence 4451, App
725	14.4	80.0	397658	9	US-09-813-320-3	Sequence 3, App	798	13.8	76.7	339	18	US-10-425-115-16293	Sequence 16293, A
726	14.4	80.0	513509	10	US-09-754-853A-4	Sequence 4, App	799	13.8	76.7	330	18	US-10-425-115-81314	Sequence 81314, A
727	14.4	80.0	401	16	US-10-424-599-29057	Sequence 29057, A	800	13.8	76.7	336	9	US-09-815-242-4180	Sequence 4180, App
728	14.4	80.0	462	17	US-10-437-963-71009	Sequence 71009, A	801	13.8	76.7	351	16	US-10-282-122A-4362	Sequence 362, A
729	14.4	80.0	503	10	US-09-770-961-359	Sequence 359, App	802	13.8	76.7	354	9	US-09-815-242-8160	Sequence 8160, App
730	14.4	80.0	539	17	US-10-767-701-2505	Sequence 2505, App	803	13.8	76.7	354	9	US-10-282-122A-7931	Sequence 7931, App
731	14.4	80.0	552	18	US-10-653-047-603	Sequence 603, App	804	13.8	76.7	354	16	US-10-424-599-136239	Sequence 136239, A
732	14.4	80.0	577	18	US-10-425-115-15364	Sequence 15364, A	805	13.8	76.7	370	18	US-10-674-124A-53239	Sequence 53239, A
733	14.4	80.0	627	15	US-10-369-493-31155	Sequence 31155, A	806	13.8	76.7	394	16	US-10-424-599-116618	Sequence 116618, A
734	14.4	80.0	637	15	US-10-369-493-28396	Sequence 28396, A	807	13.8	76.7	406	18	US-10-425-115-119584	Sequence 119584, A
735	14.4	80.0	669	13	US-10-027-632-203812	Sequence 203812, A	808	13.8	76.7	411	9	US-09-834-975-155	Sequence 156, App
736	14.4	80.0	669	13	US-10-027-632-203813	Sequence 203813, A	809	13.8	76.7	435	16	US-10-424-599-81087	Sequence 81087, A
737	14.4	80.0	669	15	US-10-027-632-203812	Sequence 203812, A	810	13.8	76.7	435	17	US-10-437-963-27667	Sequence 27667, A
738	14.4	80.0	669	15	US-10-027-632-203813	Sequence 203813, A	811	13.8	76.7	435	18	US-10-674-124A-53239	Sequence 53239, A
739	14.4	80.0	676	18	US-10-653-047-4405	Sequence 4405, App	812	13.8	76.7	454	18	US-10-425-115-105070	Sequence 105070, A
740	14.4	80.0	735	16	US-10-283-122A-31241	Sequence 31241, A	813	13.8	76.7	452	9	US-09-070-927A-962	Sequence 962, App
741	14.4	80.0	835	16	US-10-424-599-99215	Sequence 99215, A	814	13.8	76.7	455	9	US-09-864-761-2597	Sequence 2597, App
742	14.4	80.0	857	9	US-09-770-445-609	Sequence 609, App	815	13.8	76.7	469	18	US-10-425-115-28626	Sequence 28626, A

816	13.8	76.7	469	18	US-10-425-115-176565	Sequence 176565,
817	13.8	76.7	478	9	US-09-908-931B-47	Sequence 47, Appl
818	13.8	76.7	478	16	US-10-333-964-48	Sequence 48, Appl
819	13.8	76.7	480	16	US-10-260-238-5147	Sequence 5147, Ap
820	13.8	76.7	481	16	US-10-424-559-41021	Sequence 41021, A
821	13.8	76.7	486	16	US-10-260-238-1471	Sequence 1471, Ap
822	13.8	76.7	506	16	US-10-424-559-139527	Sequence 139527, A
823	13.8	76.7	509	16	US-10-437-963-57351	Sequence 57351, A
824	13.8	76.7	512	15	US-10-029-386-6316	Sequence 92836, Ap
825	13.8	76.7	519	15	US-10-027-632-92836	Sequence 92836, A
826	13.8	76.7	519	15	US-10-027-632-92836	Sequence 14778, A
827	13.8	76.7	520	18	US-10-425-115-14778	Sequence 2908, Ap
828	13.8	76.7	529	17	US-10-767-701-2908	Sequence 822, App
829	13.8	76.7	531	9	US-09-770-152-822	Sequence 114152,
830	13.8	76.7	545	16	US-10-424-559-114152	Sequence 80, Appl
831	13.8	76.7	546	15	US-10-349-680-80	Sequence 31358, A
832	13.8	76.7	558	16	US-10-425-114-31358	Sequence 46, Appl
833	13.8	76.7	580	15	US-10-161-051-46	Sequence 89488, A
834	13.8	76.7	586	17	US-10-437-963-89488	Sequence 152937,
835	13.8	76.7	588	18	US-10-425-115-152937	Sequence 66786, A
836	13.8	76.7	589	13	US-10-027-632-66786	Sequence 67330, A
837	13.8	76.7	589	13	US-10-027-632-67330	Sequence 311242,
838	13.8	76.7	589	13	US-10-027-632-311242	Sequence 66786, A
839	13.8	76.7	589	15	US-10-027-632-66786	Sequence 67330, A
840	13.8	76.7	589	15	US-10-027-632-67330	Sequence 311242,
841	13.8	76.7	589	15	US-10-027-632-311242	Sequence 311242,
842	13.8	76.7	590	17	US-10-021-323-311242	Sequence 7885, Ap
843	13.8	76.7	590	17	US-10-021-323-7885	Sequence 6950, Ap
844	13.8	76.7	605	18	US-10-425-115-6950	Sequence 6548, Ap
845	13.8	76.7	619	14	US-09-796-692-6548	Sequence 6548, Ap
846	13.8	76.7	619	14	US-10-050-863-6548	Sequence 6548, Ap
847	13.8	76.7	619	16	US-10-047-478B-6548	Sequence 6548, Ap
848	13.8	76.7	619	16	US-10-154-884B-6548	Sequence 6548, Ap
849	13.8	76.7	619	17	US-10-764-324-6548	Sequence 1433, Ap
850	13.8	76.7	634	17	US-10-767-701-1433	Sequence 9002, Ap
851	13.8	76.7	634	13	US-10-198-846-9002	Sequence 209173,
852	13.8	76.7	637	13	US-10-027-632-209173	Sequence 209173,
853	13.8	76.7	637	15	US-10-027-632-209173	Sequence 277, App
854	13.8	76.7	640	16	US-10-425-114-279	Sequence 16159, A
855	13.8	76.7	645	18	US-10-363-345A-16159	Sequence 16160, A
856	13.8	76.7	645	18	US-10-363-345A-16160	Sequence 1748, Ap
857	13.8	76.7	653	9	US-09-974-300-1748	Sequence 203055,
858	13.8	76.7	676	13	US-10-027-632-203055	Sequence 6182, Ap
859	13.8	76.7	676	15	US-10-027-632-203055	Sequence 6182, Ap
860	13.8	76.7	676	18	US-10-653-047-6182	Sequence 170015, Ap
861	13.8	76.7	687	18	US-10-425-115-6852	Sequence 12932, A
862	13.8	76.7	695	18	US-10-425-115-12932	Sequence 6071, Ap
863	13.8	76.7	696	18	US-10-363-345A-6071	Sequence 6072, Ap
864	13.8	76.7	705	9	US-10-363-345A-6072	Sequence 1824, Ap
865	13.8	76.7	720	18	US-09-974-300-1824	Sequence 183719,
866	13.8	76.7	728	18	US-10-425-115-183719	Sequence 183719,
867	13.8	76.7	728	18	US-10-012-697-1844	Sequence 29124, Ap
868	13.8	76.7	734	17	US-10-437-963-29124	Sequence 1318, Ap
869	13.8	76.7	734	18	US-10-739-930-1318	Sequence 183723, A
870	13.8	76.7	751	18	US-10-425-115-183723	Sequence 68522, A
871	13.8	76.7	753	17	US-10-437-963-68522	Sequence 3709, Ap
872	13.8	76.7	759	9	US-09-764-877-3709	Sequence 3709, Ap
873	13.8	76.7	769	16	US-10-242-515-3709	Sequence 8011, Ap
874	13.8	76.7	773	16	US-09-845-442-8012	Sequence 30650, A
875	13.8	76.7	783	16	US-10-282-122A-30660	Sequence 50, Appl
876	13.8	76.7	801	13	US-09-841-260-50	Sequence 50, Appl
877	13.8	76.7	801	13	US-10-007-693-50	Sequence 50, Appl
878	13.8	76.7	801	17	US-10-762-058-50	Sequence 183718,
879	13.8	76.7	805	18	US-10-425-115-183718	Sequence 183744,
880	13.8	76.7	823	16	US-10-424-559-137644	Sequence 6642, Ap
881	13.8	76.7	846	16	US-10-382-122A-6642	Sequence 9576, Ap
882	13.8	76.7	848	14	US-10-198-846-9576	Sequence 116, App
883	13.8	76.7	849	15	US-10-430-011-116	Sequence 36688, A
884	13.8	76.7	849	16	US-10-282-122A-36688	Sequence 39089, A
885	13.8	76.7	852	16	US-10-282-122A-39089	Sequence 39969, A
886	13.8	76.7	852	16	US-10-282-122A-39969	Sequence 15497, A
887	13.8	76.7	856	16	US-10-424-559-15497	Sequence 399, App
888	13.8	76.7	869	9	US-09-071-035-399	Sequence 399, App

```
962 13.8 76.7 1599 15 US-10-128-714-2442 Sequence 2442, Ap
963 13.8 76.7 1599 15 US-10-128-714-2442 Sequence 2442, Ap
964 13.8 76.7 1614 16 US-10-425-114-14594 Sequence 14594, A
965 13.8 76.7 1639 18 US-10-425-115-56500 Sequence 56500, A
966 13.8 76.7 1643 16 US-10-425-114-26886 Sequence 26886, A
967 13.8 76.7 1643 18 US-10-425-115-75255 Sequence 75255, A
968 13.8 76.7 1670 18 US-10-425-115-178309 Sequence 178309,
969 13.8 76.7 1680 9 US-09-071-035-85 Sequence 85, Appl
970 13.8 76.7 1680 14 US-10-206-576-85 Sequence 85, Appl
971 13.8 76.7 1705 16 US-10-425-114-6395 Sequence 6395, Ap
972 13.8 76.7 1716 15 US-10-369-493-34585 Sequence 34585, A
973 13.8 76.7 1726 17 US-10-767-701-13885 Sequence 13885, A
974 13.8 76.7 1739 15 US-10-128-714-1442 Sequence 1442, Ap
975 13.8 76.7 1739 15 US-10-128-714-6442 Sequence 6442, Ap
976 13.8 76.7 1749 9 US-09-938-842A-3369 Sequence 3369, Ap
977 13.8 76.7 1749 11 US-09-938-842A-3369 Sequence 3369, Ap
978 13.8 76.7 1760 18 US-10-425-115-90240 Sequence 90240, A
979 13.8 76.7 1762 18 US-10-425-115-178311 Sequence 178311,
980 13.8 76.7 1819 14 US-10-219-220-243 Sequence 243, Appl
981 13.8 76.7 1829 9 US-09-764-864-190 Sequence 190, Appl
982 13.8 76.7 1829 18 US-10-739-930-2921 Sequence 2921, Ap
983 13.8 76.7 1833 16 US-10-424-599-90755 Sequence 90755, A
984 13.8 76.7 1834 16 US-10-424-599-90755 Sequence 90755, A
985 13.8 76.7 1846 18 US-10-739-930-4328 Sequence 4328, Ap
986 13.8 76.7 1854 15 US-10-037-270-793 Sequence 793, Appl
987 13.8 76.7 1854 15 US-10-117-722-793 Sequence 793, Appl
988 13.8 76.7 1874 10 US-09-822-846-188 Sequence 188, Appl
989 13.8 76.7 1883 16 US-10-374-780A-1759 Sequence 1759, Ap
990 13.8 76.7 1922 18 US-10-425-115-178313 Sequence 178313,
991 13.8 76.7 1927 18 US-10-425-115-178957 Sequence 178957,
992 13.8 76.7 1949 17 US-10-115-635-55 Sequence 55, Appl
993 13.8 76.7 1965 16 US-10-282-122A-7372 Sequence 7372, Ap
994 13.8 76.7 1995 18 US-10-425-115-56269 Sequence 56269, A
995 13.8 76.7 2000 9 US-09-887-576-147 Sequence 147, Appl
996 13.8 76.7 2000 9 US-09-938-842A-4847 Sequence 4847, Ap
997 13.8 76.7 2000 11 US-09-938-842A-4847 Sequence 4847, Ap
998 13.8 76.7 2027 9 US-09-771-161A-29 Sequence 29, Appl
999 13.8 76.7 2055 14 US-10-004-633-3 Sequence 3, Appl
1000 13.8 76.7 2055 15 US-10-252-131-3 Sequence 3, Appl
```

ALIGNMENTS

```
RESULT 1
US-10-050-189a-6
; Sequence 6, Application US/10050189A
; Publication No. US2002016856A1
; GENERAL INFORMATION:
; APPLICANT: Rubin, Berish
; TITLE OF INVENTION: Detection of Mutations in a Gene Encoding IKB Kinase-Complex-Asso
; FILE REFERENCE: Rubin 201
; CURRENT APPLICATION NUMBER: US/10/050,189A
; CURRENT FILING DATE: 2002-01-16
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 18
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-050-189A-6
```

```
Query Match 100.0%; Score 18; DB 13; Length 18;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 1 GAGAACACACAGATTTC 18
Db 1 GAGAACACACAGATTTC 18
```

```
RESULT 2
US-10-027-632-111967/c
; Sequence 111967, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108927.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 111967
; LENGTH: 2465
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-111967
```

```
Query Match 88.9%; Score 16; DB 13; Length 2465;
Best Local Similarity 100.0%; Pred. No. 56+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 1 GAGAACACACAGATTTC 16
Db 1560 GAGAACACACAGATTTC 1545
```

```
RESULT 3
US-10-027-632-111967/c
; Sequence 111967, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108927.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 111967
; LENGTH: 2465
; TYPE: DNA
; ORGANISM: Human
```


US-10-027-632-111967

Query Match 88.9%; Score 16; DB 15; Length 2465;
Best Local Similarity 100.0%; Pred. No. 5e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAGACACACAGATTTC 16
DB 1560 GAGACACACAGATTTC 1545

RESULT 4
US-10-041-856-1

; Sequence 1, Application US/10041856
; Publication No. US20020169299A1
; GENERAL INFORMATION:
; APPLICANT: SLAUGENHAPT, SUSAN
; APPLICANT: GUSELLA, JAMES P.
; TITLE OF INVENTION: GENE FOR IDENTIFYING INDIVIDUALS WITH FAMILIAL
; TITLE OF INVENTION: DYSAUTONOMIA
; FILE REFERENCE: 1829-4004US1
; CURRENT APPLICATION NUMBER: US/10/041,856
; CURRENT FILING DATE: 2002-07-08
; PRIOR APPLICATION NUMBER: 60/260,080
; PRIOR FILING DATE: 2001-01-06
; NUMBER OF SEQ ID NOS: 88
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 66479
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-041-856-1

Query Match 88.9%; Score 16; DB 13; Length 66479;
Best Local Similarity 100.0%; Pred. No. 7.2e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAGACACACAGATTTC 16
DB 33979 GAGACACACAGATTTC 33994

RESULT 5

US-09-833-381-908
; Sequence 908, Application US/09833381
; Patent No. US20020132090A1
; GENERAL INFORMATION:
; APPLICANT: ROBISON, Keith E.
; TITLE OF INVENTION: No. US20020132090A1el Nucleic Acid and Protein Homologs
; FILE REFERENCE: 5800-119
; CURRENT APPLICATION NUMBER: US/09/833,381
; CURRENT FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: 09/516,448
; PRIOR FILING DATE: 2000-02-29
; NUMBER OF SEQ ID NOS: 2050
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 908
; LENGTH: 409
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(409)
; OTHER INFORMATION: n = A,T,C or G
US-09-833-381-908

Query Match 85.6%; Score 15.4; DB 9; Length 409;
Best Local Similarity 94.1%; Pred. No. 8.1e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 AGACACACAGATTTCG 18
DB 170 AGACACACAGATTTCG 186

RESULT 6
US-09-878-178-1155

; Sequence 1155, Application US/09878178
; Patent No. US20020177552A1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yugu
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Secrist, Heather
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF COLON CANCER
; FILE REFERENCE: 21021.527
; CURRENT APPLICATION NUMBER: US/09/878,178
; CURRENT FILING DATE: 2001-06-08
; NUMBER OF SEQ ID NOS: 2237
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1155
; LENGTH: 482
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-878-178-1155

Query Match 85.6%; Score 15.4; DB 9; Length 482;
Best Local Similarity 94.1%; Pred. No. 8.3e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 AGACACACAGATTTCG 18
DB 48 AGACACACAGATTTCG 64

RESULT 7

US-10-046-935-1155
; Sequence 1155, Application US/10046935
; Publication No. US20020156011A1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yugu
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Secrist, Heather
; APPLICANT: Wang, Aijun
; APPLICANT: Stolk, John A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF COLON CANCER
; FILE REFERENCE: 21021.527C1
; CURRENT APPLICATION NUMBER: US/10/046,935
; CURRENT FILING DATE: 2002-01-15
; NUMBER OF SEQ ID NOS: 2239
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1155
; LENGTH: 482
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-046-935-1155

Query Match 85.6%; Score 15.4; DB 13; Length 482;
Best Local Similarity 94.1%; Pred. No. 8.3e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 AGACACACAGATTTCG 18
DB 48 AGACACACAGATTTCG 64

RESULT 8

US-10-146-502-1155
; Sequence 1155, Application US/10146502
; Publication No. US20030069180A1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yugu
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Secrist, Heather
; APPLICANT: Wang, Aijun

```

; APPLICANT: Stolk, John A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.527C2
; CURRENT APPLICATION NUMBER: US/10/146,502
; CURRENT FILING DATE: 2002-05-14
; NUMBER OF SEQ ID NOS: 2241
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1155
; LENGTH: 462
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-146-502-1155

Query Match      85.6%; Score 15.4; DB 14; Length 482;
Best Local Similarity 94.1%; Pred. No. 8.3e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      2 AGAACACAAGATTCCG 18
         |||||
Db      48 AGCACACAAGATTCCG 64

RESULT 9
US-10-282-122A-20772
; Sequence 20772, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: EITPA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/151,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 20772
; TYPE: DNA
; ORGANISM: Enterococcus faecalis
US-10-282-122A-20772
```

```

Query Match      85.6%; Score 15.4; DB 16; Length 513;
Best Local Similarity 94.1%; Pred. No. 8.3e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      2 AGAACACAAGATTCCG 18
         |||||
Db      447 AGAACACAAGATTCCG 463

RESULT 10
US-10-027-632-277903/c
; Sequence 277903, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,493
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 277903
; LENGTH: 577
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-277903

Query Match      85.6%; Score 15.4; DB 13; Length 577;
Best Local Similarity 94.1%; Pred. No. 8.5e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      2 AGAACACAAGATTCCG 18
         |||||
Db      31 AGCACACAAGATTCCG 15

RESULT 11
US-10-027-632-277903/c
; Sequence 277903, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
```

;; PRIOR FILING DATE: 1999-09-28
;; PRIOR APPLICATION NUMBER: US 60/146,002
;; PRIOR FILING DATE: 1999-08-09
;; NUMBER OF SEQ ID NOS: 325720
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 277903
;; LENGTH: 577
;; TYPE: DNA
;; ORGANISM: Human
US-10-027-632-277903

Query Match 85.6%; Score 15.4; DB 15; Length 577;
Best Local Similarity 94.1%; Pred. No. 8.5e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 AGAACACACAGATTGCG 18
|||
Db 31 AGCACACACAGATTGCG 15

RESULT 12
US-10-027-632-277719
; Sequence 277719, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/199,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 277719
; LENGTH: 582
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-277719

Query Match 85.6%; Score 15.4; DB 13; Length 582;
Best Local Similarity 94.1%; Pred. No. 8.5e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 AGAACACACAGATTGCG 18
|||
Db 553 AGCACACACAGATTGCG 569

RESULT 13
US-10-027-632-277719
; Sequence 277719, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30

;; PRIOR APPLICATION NUMBER: US 60/218,006
;; PRIOR FILING DATE: 2000-07-12
;; PRIOR APPLICATION NUMBER: US 60/198,676
;; PRIOR FILING DATE: 2000-04-20
;; PRIOR APPLICATION NUMBER: US 60/193,483
;; PRIOR FILING DATE: 2000-03-29
;; PRIOR APPLICATION NUMBER: US 60/185,218
;; PRIOR FILING DATE: 2000-02-24
;; PRIOR APPLICATION NUMBER: US 60/167,363
;; PRIOR FILING DATE: 1999-11-23
;; PRIOR APPLICATION NUMBER: US 60/156,358
;; PRIOR FILING DATE: 1999-09-28
;; PRIOR APPLICATION NUMBER: US 60/146,002
;; PRIOR FILING DATE: 1999-08-09
;; NUMBER OF SEQ ID NOS: 325720
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 277719
;; LENGTH: 582
;; TYPE: DNA
;; ORGANISM: Human
US-10-027-632-277719

Query Match 85.6%; Score 15.4; DB 15; Length 582;
Best Local Similarity 94.1%; Pred. No. 8.5e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 AGAACACACAGATTGCG 18
|||
Db 553 AGCACACACAGATTGCG 569

RESULT 14
US-10-767-701-3256/c
; Sequence 3256, Application US/10767701
; Publication No. US20040172684A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement
; FILE REFERENCE: 38-21(5355)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 3256
; LENGTH: 582
; TYPE: DNA
; ORGANISM: Sorghum bicolor
; FEATURE:
; OTHER INFORMATION: Clone ID: SORBI-28MAY03-CLUS55082_1
US-10-767-701-3256

Query Match 85.6%; Score 15.4; DB 17; Length 582;
Best Local Similarity 94.1%; Pred. No. 8.5e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 AGAACACACAGATTGCG 18
|||
Db 406 AGAACACACATATTGCG 390

RESULT 15
US-10-437-963-32919/c
; Sequence 32919, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.

```

; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Nice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 32919
; LENGTH: 775
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_37082C.1
US-10-437-963-32919

Query Match      85.6%; Score 15.4; DB 17; Length 775;
Best Local Similarity 94.1%; Pred. No. 8.7e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      2 AGAACACACAGATTCCG 18
Db      598 AGAAGTACACAGATTCCG 582

RESULT 16
US-09-815-242-6118/c
; Sequence 6118, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6118
; LENGTH: 1470
; TYPE: DNA
; ORGANISM: Escherichia coli
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1) ... (1470)
US-09-815-242-6118

Query Match      85.6%; Score 15.4; DB 9; Length 1470;
Best Local Similarity 94.1%; Pred. No. 9.4e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      2 AGAACACACAGATTCCG 18

```

```

Db      757 AGAACACACAGATTCCG 741

RESULT 17
US-10-369-493-24585/c
; Sequence 24585, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Cao, Yongwei
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 24585
; LENGTH: 1470
; TYPE: DNA
; ORGANISM: Escherichia coli
US-10-369-493-24585

Query Match      85.6%; Score 15.4; DB 15; Length 1470;
Best Local Similarity 94.1%; Pred. No. 9.4e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      2 AGAACACACAGATTCCG 18
Db      757 AGAACACACAGATTCCG 741

RESULT 18
US-10-369-493-32483
; Sequence 32483, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 32483
; LENGTH: 1518
; TYPE: DNA
; ORGANISM: Ralstonia metallidurans
US-10-369-493-32483

Query Match      85.6%; Score 15.4; DB 15; Length 1518;
Best Local Similarity 94.1%; Pred. No. 9.4e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      2 AGAACACACAGATTCCG 18
Db      1043 AGTACACACAGATTCCG 1059

RESULT 19
US-10-369-493-45765
; Sequence 45765, Application US/10369493

```

```
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10 (52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; PRIOR FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 45765
; LENGTH: 1950
; TYPE: DNA
; ORGANISM: Saccharomyces cerevisiae
US-10-369-493-45765

Query Match      85.6%; Score 15.4; DB 15; Length 1950;
Best Local Similarity 94.1%; Pred. No. 9.7e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY      2 AGAACAACAAGATTGCG 18
Db      1795 AGAAGACAAGATTGCG 1811

RESULT 20
US-10-169-257-8/c
; Sequence 8, Application US/10169257
; Publication No. US20030207410A1
; GENERAL INFORMATION:
; APPLICANT: Tadashi FUJII
; APPLICANT: Yasuhide ARITOKU
; APPLICANT: Manabu MUKAIHARA
; APPLICANT: Takao NARITA
; APPLICANT: Hitosi AGEMATU
; APPLICANT: Kunio ISSHITI
; TITLE OF INVENTION: Biological process for the production of L-pipecolic acid
; FILE REFERENCE: 2002-0845A/WMC/00202
; CURRENT APPLICATION NUMBER: US/10/169,257
; PRIOR FILING DATE: 2002-06-28
; PRIOR APPLICATION NUMBER: JP11/373389
; PRIOR FILING DATE: 1999-12-28
; NUMBER OF SEQ ID NOS: 20
; SEQ ID NO 8
; LENGTH: 2186
; TYPE: DNA
; ORGANISM: Escherichia coli
; FEATURE:
US-10-169-257-8

Query Match      85.6%; Score 15.4; DB 15; Length 2186;
Best Local Similarity 94.1%; Pred. No. 9.8e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY      2 AGAACAACAAGATTGCG 18
Db      1281 AGAACAACAAGATTGCG 1265

RESULT 21
US-09-308-207-7
; Sequence 7, Application US/09308207
; Publication No. US20030022323A1
; GENERAL INFORMATION:
; APPLICANT: MARIA DIAZ-TORRES ET AL.
; TITLE OF INVENTION: METHOD FOR THE RECOMBINANT
; PRODUCTION OF 1,3 PROPAEDIOL
; NUMBER OF SEQUENCES: 68
```

```
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genencor International, Inc.
; STREET: 4 Cambridge Place
; CITY: Rochester
; STATE: NY
; COUNTRY: U.S.A
; ZIP: 14618
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows
; SOFTWARE: FastSeq for Windows Version 2.0b
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/308,207
; FILING DATE: 13-May-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/030,601
; FILING DATE: 13-NOV-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Glaister, Debra
; REGISTRATION NUMBER: 33,888
; REFERENCE/DOCKET NUMBER: GC 369-2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-864-7620
; TELEFAX: 650-845-6504
; Telex: <Unknown>
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3178 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; ORGANISM: GUT2
; SEQUENCE DESCRIPTION: SEQ ID NO: 7:
US-09-308-207-7

Query Match      85.6%; Score 15.4; DB 10; Length 3178;
Best Local Similarity 94.1%; Pred. No. 1e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY      2 AGAACAACAAGATTGCG 18
Db      2154 AGAACAACAAGATTGCG 2170

RESULT 22
US-10-093-463-135
; Sequence 135, Application US/10093463
; Publication No. US20030208039A1
; GENERAL INFORMATION:
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Shemoy, Suresh
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Gusev, Vladimir
; APPLICANT: Pochart, Pascal
; APPLICANT: Zhong, Mei
; APPLICANT: Rastelli, Luca
; APPLICANT: Mezes, Peter
; APPLICANT: Smithson, Glenda
; APPLICANT: Guo, Xiaojia
; APPLICANT: Gerlach, Valerie
; APPLICANT: Casman, Stacie
; APPLICANT: Boldog, Ferenc
; APPLICANT: Li, Li
; APPLICANT: Zethusen, Bryan
; APPLICANT: Tchernev, Velizar
; APPLICANT: Gangoli, Esna
; APPLICANT: Verne, Corine
; APPLICANT: Pena, Carol
```

```

APPLICANT: Burgess, Catherine
APPLICANT: Liu, Xiaohong
APPLICANT: Spytek, Kimberly
APPLICANT: Gorman, Linda
APPLICANT: Spaderna, Steven
APPLICANT: Voss, Edward
APPLICANT: Malyankar, Uriel
APPLICANT: Anderson, David
APPLICANT: Patturajan, Meera
APPLICANT: Miller, Charles
APPLICANT: Taupier, Raymond J. Jr.
TITLE OF INVENTION: No. US20030208039A1el Antibodies that Bind to Antigenic Polypeptide
FILE REFERENCE: 21402-290A (Cura 590A7)
CURRENT FILING DATE: 2002-06-24
PRIOR APPLICATION NUMBER: 60/283,675
PRIOR FILING DATE: 2001-04-14
PRIOR APPLICATION NUMBER: 60/338,092
PRIOR FILING DATE: 2001-12-03
PRIOR APPLICATION NUMBER: 60/274,281
PRIOR FILING DATE: 2001-03-08
PRIOR APPLICATION NUMBER: 60/274,101
PRIOR FILING DATE: 2001-03-08
PRIOR APPLICATION NUMBER: 60/325,681
PRIOR FILING DATE: 2001-09-27
PRIOR APPLICATION NUMBER: 60/304,354
PRIOR FILING DATE: 2001-07-10
PRIOR APPLICATION NUMBER: 60/279,995
PRIOR FILING DATE: 2001-03-30
PRIOR APPLICATION NUMBER: 60/294,899
PRIOR FILING DATE: 2001-05-31
PRIOR APPLICATION NUMBER: 60/287,424
PRIOR FILING DATE: 2001-04-30
PRIOR APPLICATION NUMBER: 60/299,027
PRIOR FILING DATE: 2001-06-18
PRIOR APPLICATION NUMBER: 60/309,198
PRIOR FILING DATE: 2001-07-31
PRIOR APPLICATION NUMBER: 60/281,194
PRIOR FILING DATE: 2001-04-04
PRIOR APPLICATION NUMBER: 60/274,194
PRIOR FILING DATE: 2001-03-08
PRIOR APPLICATION NUMBER: 60/274,849
PRIOR FILING DATE: 2001-03-09
PRIOR APPLICATION NUMBER: 60/330,380
PRIOR FILING DATE: 2001-10-18
PRIOR APPLICATION NUMBER: 60/275,235
PRIOR FILING DATE: 2001-03-12
PRIOR APPLICATION NUMBER: 60/288,342
PRIOR FILING DATE: 2001-05-03
PRIOR APPLICATION NUMBER: 60/275,578
PRIOR FILING DATE: 2001-03-13
NUMBER OF SEQ ID NOS: 370
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 135
LENGTH: 5101
TYPE: DNA
ORGANISM: Homo sapiens
FEATURES:
NAME/KEY: CDS
LOCATION: (77)..(3308)
US-10-093-463-135
Query Match 85.6%; Score 15.4; DB 15; Length 5101;
Best Local Similarity 94.1%; Pred. No. 1,1e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
2 AGACACACAGATTGCG 18
396 AGCACACAGATTGCG 412

```

```

US-10-367-094-100
; Sequence 100, Application US/10367094
; Publication No. US20040170982a1
GENERAL INFORMATION:
; APPLICANT: David W. Morris
; APPLICANT: Marc Malandro
; TITLE OF INVENTION: Novel Therapeutic Targets in Cancer
; FILE REFERENCE: 529452001500
; CURRENT APPLICATION NUMBER: US/10/367,094
; CURRENT FILING DATE: 2003-02-14
; NUMBER OF SEQ ID NOS: 203
; SOFTWARE: PastSeq for Windows Version 4.0
; SEQ ID NO 100
; LENGTH: 5719
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-367-094-100

Query Match      85.6%; Score 15.4; DB 17; Length 5719;
Best Local Similarity 94.1%; Pred.No. 1.le+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Cy      2 AGAACACACAGATTCCG 18
      |||||||
Db      368 AGCACACACAGATTCCG 364

RESULT 24
US-09-070-927A-288
; Sequence 288, Application US/09070927A
; Patent No. US20020120116a1
GENERAL INFORMATION:
; APPLICANT: Charles A. Kunsch
; APPLICANT: Patrick J. Dillon
; Steven Barash
; TITLE OF INVENTION: Enterococcus faecialis Polynucleotides and Polypeptides
; NUMBER OF SEQUENCES: 982
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/070,927A
; FILING DATE: 04-May-2000
; CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/046,655
; FILING DATE: 1997-05-16
; APPLICATION NUMBER: 60/044,031
; FILING DATE: 1997-05-06
; APPLICATION NUMBER: 60/066,009
; FILING DATE: 1997-11-14
ATTORNEY/AGENT INFORMATION:
; NAME: Kenley K. Hoover
; REGISTRATION NUMBER: 40,302
; REFERENCE/DOCKET NUMBER: P3269
TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 288:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7043 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear

```

SEQUENCE DESCRIPTION: SEQ ID NO: 268;
US-09-070-927a-268

Query Match 85.6%; Score 15.4; DB 9; Length 7043;
Best Local Similarity 94.1%; Pred. No. 1.1e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 AGACACACAGATTCCG 18
DB 1018 AGACACACAGATTCCG 1034

RESULT 25
US-10-367-094-99
; Sequence 99, Application US/10367094
; Publication No. US20040170982A1
; GENERAL INFORMATION:
; APPLICANT: David W. Morris
; APPLICANT: Marc Malandro
; TITLE OF INVENTION: Novel Therapeutic Targets in Cancer
; FILE REFERENCE: 529452001500
; CURRENT APPLICATION NUMBER: US/10/367,094
; CURRENT FILING DATE: 2003-02-14
; NUMBER OF SEQ ID NOS: 203
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 99
; LENGTH: 142976
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(142976)
; OTHER INFORMATION: n = A,T,C or G
US-10-367-094-99

Query Match 85.6%; Score 15.4; DB 17; Length 142976;
Best Local Similarity 94.1%; Pred. No. 1.5e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 AGACACACAGATTCCG 18
DB 93623 AGACACACAGATTCCG 93639

RESULT 26
US-10-437-963-86110
; Sequence 86110, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 86110
; LENGTH: 378
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_85182C.1
US-10-437-963-86110

Query Match 83.3%; Score 15; DB 17; Length 378;
Best Local Similarity 100.0%; Pred. No. 1.3e+03;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAGACACACAGATT 15
DB 362 GAGACACACAGATT 376

RESULT 27
US-10-137-036-4/C
; Sequence 4, Application US/10137036
; Publication No. US20030101478A1
; GENERAL INFORMATION:
; APPLICANT: Pereira, Ranjan
; APPLICANT: Rice, Stephen
; APPLICANT: Eagleton, Clare
; APPLICANT: Laaham, Annette
; APPLICANT: Wood, Marion
; APPLICANT: Visser, Elizabeth
; TITLE OF INVENTION: Compositions and Methods for the
; TITLE OF INVENTION: Modification of Gene Expression
; FILE REFERENCE: 11000.1036c4
; CURRENT APPLICATION NUMBER: US/10/137,036
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: PCT/02/01115
; PRIOR FILING DATE: 2001-06-20
; PRIOR APPLICATION NUMBER: U.S. No. US20030101478A1 09/724,624
; PRIOR FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: U.S. No. US20030101478A1 09/598,401
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: PCT/NZ00/00018
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: U.S. No. US20030101478A1 60/146,591
; PRIOR FILING DATE: 1999-07-30
; PRIOR APPLICATION NUMBER: U.S. No. US20030101478A1 09/276,599
; NUMBER OF SEQ ID NOS: 143
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 485
; TYPE: DNA
; ORGANISM: Pinus radiata
; FEATURE:
; NAME/KEY: 5'UTR
; LOCATION: (1)...(431)
; FEATURE:
; NAME/KEY: TATA_signal
; LOCATION: (350)...(356)
; FEATURE:
; NAME/KEY: CAAT_signal
; LOCATION: (326)...(333)
US-10-137-036-4

Query Match 83.3%; Score 15; DB 15; Length 485;
Best Local Similarity 100.0%; Pred. No. 1.3e+03;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 AACACACAGATTCCG 18
DB 409 AACACACAGATTCCG 395

RESULT 28
US-09-770-149-77/C
; Sequence 77, Application US/09770149
; Patent No. US2002005963A1
; GENERAL INFORMATION:
; APPLICANT: Gorlach, Jörn
; APPLICANT: An, Yong-Qiang
; APPLICANT: Hamilton, Carol M.
; APPLICANT: Price, Jennifer L.
; APPLICANT: Raines, Tracy M.
; APPLICANT: Yu, Yang
; APPLICANT: Rameaka, Joshua G.

APPLICANT: Page, Amy
APPLICANT: Mathew, Abraham V.
APPLICANT: Ledford, Brooke L.
APPLICANT: Woessner, Jeffrey P.
APPLICANT: Haas, William David
APPLICANT: Garcia, Carlos A.
APPLICANT: Krickler, Maja
APPLICANT: Slader, Ted
APPLICANT: Davis, Keith R.
APPLICANT: Allen, Keith
APPLICANT: Hoffman, Neil
APPLICANT: Hurban, Patrick
TITLE OF INVENTION: Expressed Sequences of Arabidopsis
TITLE OF INVENTION: thaliana
FILE REFERENCE: 2024 (PARA-013PRV)
CURRENT APPLICATION NUMBER: US/09/770,149
CURRENT FILING DATE: 2001-01-26
PRIORITY FILING DATE: 2000-01-27
NUMBER OF SEQ ID NOS: 999
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 77
LENGTH: 737
TYPE: DNA
ORGANISM: Arabidopsis thaliana
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)..(737)
OTHER INFORMATION: n = A,T,C or G
US-09-770-149-77

Query Match 83.3%; Score 15; DB 9; Length 737;
Best Local Similarity 100.0%; Pred. No. 1.4e+03;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CY 1 GAGACACACAGATT 15
DB 667 GAGACACACAGATT 653

RESULT 29
US-10-767-701-11437/c
Sequence 11437, Application US/10767701
Publication No. US20040172684A1
GENERAL INFORMATION:
APPLICANT: Kovalic, David K.
APPLICANT: Zhou, Yihua
APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
FILE REFERENCE: 38-21(53535)B
CURRENT APPLICATION NUMBER: US/10/767,701
CURRENT FILING DATE: 2004-01-29
NUMBER OF SEQ ID NOS: 63128
SEQ ID NO 11437
LENGTH: 1169
TYPE: DNA
ORGANISM: Sorghum bicolor
FEATURE:
OTHER INFORMATION: Clone ID: SORBI-28MAY03-CLUS405_1
US-10-767-701-11437

Query Match 83.3%; Score 15; DB 17; Length 1169;
Best Local Similarity 100.0%; Pred. No. 1.4e+03;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CY 2 AGAACACACAGATT 16
DB 309 AGAACACACAGATT 295

RESULT 30
US-10-032-585-6133

Sequence 6133, Application US/10032585
Publication No. US20030180953A1
GENERAL INFORMATION:
APPLICANT: Terry, Roemer D.
APPLICANT: Bo, Jiang
APPLICANT: Charles, Boone
APPLICANT: Howard, Busssey
TITLE OF INVENTION: Gene Disruption Methodologies for Drug Target Discovery
FILE REFERENCE: 10182-005-999
CURRENT APPLICATION NUMBER: US/10/032,585
CURRENT FILING DATE: 2001-12-20
NUMBER OF SEQ ID NOS: 8000
SOFTWARE: PatentIn version 3.1
SEQ ID NO 6133
LENGTH: 1455
TYPE: DNA
ORGANISM: Candida albicans
FEATURE:
NAME/KEY: misc_feature
LOCATION: (237)..(237)
OTHER INFORMATION: n=g, a, t or c
NAME/KEY: misc_feature
LOCATION: (348)..(348)
OTHER INFORMATION: n=g, a, t or c
US-10-032-585-6133

Query Match 83.3%; Score 15; DB 15; Length 1455;
Best Local Similarity 100.0%; Pred. No. 1.5e+03;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CY 2 AGAACACACAGATT 16
DB 801 AGAACACACAGATT 815

RESULT 31
US-10-425-115-87011
Sequence 87011, Application US/10425115
Publication No. US20040214272A1
GENERAL INFORMATION:
APPLICANT: La Rosa, Thomas J.
APPLICANT: Kovalic, David K.
APPLICANT: Zhou, Yihua
APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
TITLE OF INVENTION: Plants
FILE REFERENCE: 38-21(53522)B
CURRENT APPLICATION NUMBER: US/10/425,115
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 369326
SEQ ID NO 87011
LENGTH: 209
TYPE: DNA
ORGANISM: Zea mays
FEATURE:
OTHER INFORMATION: Clone ID: MRT4577_17935C.1
US-10-425-115-87011

Query Match 82.2%; Score 14.8; DB 18; Length 209;
Best Local Similarity 88.9%; Pred. No. 1.5e+03;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

CY 1 GAGACACACAGATT 18
DB 157 GAGACACACAGATT 174

RESULT 32
US-09-867-550-1599/c
Sequence 1599, Application US/09867550
Patent No. US20020082206A1
GENERAL INFORMATION:
APPLICANT: Leach, Martin D.


```

; APPLICANT: Mehraban, Fuad,
; APPLICANT: Conley, Pamela
; APPLICANT: Law, Debbie
; APPLICANT: Topper, James
; TITLE OF INVENTION: No. US20020082206A1el Polynucleotides from Atherogenic Cells and
; FILE REFERENCE: 21402-013 (Cura-313)
; CURRENT APPLICATION NUMBER: US/09/867,550
; CURRENT FILING DATE: 2001-09-20
; PRIOR APPLICATION NUMBER: USSN 60/208,427
; PRIOR FILING DATE: 2000-05-30
; NUMBER OF SEQ ID NOS: 2123
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 1599
; LENGTH: 352
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-867-550-1599

Query Match      82.2%; Score 14.8; DB 9; Length 352;
Best Local Similarity 88.9%; Pred. No. 1.6e+03;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GAGAACACAAAGATTGCG 18
DB 20 GAAACACAAAGATTGCG 3

RESULT 33
US-10-425-115-91075/C
; Sequence 91075, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 91075
; LENGTH: 416
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_183061C.1
; US-10-425-115-91075

Query Match      82.2%; Score 14.8; DB 18; Length 416;
Best Local Similarity 88.9%; Pred. No. 1.6e+03;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GAGAACACAAAGATTGCG 18
DB 234 GAGACACAAAGATTGCG 217

RESULT 34
US-10-282-122A-33603
; Sequence 33603, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
```

```

; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 33603
; LENGTH: 438
; TYPE: DNA
; ORGANISM: Pseudomonas syringae
; US-10-282-122A-33603

Query Match      82.2%; Score 14.8; DB 16; Length 438;
Best Local Similarity 88.9%; Pred. No. 1.6e+03;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GAGAACACAAAGATTGCG 18
DB 85 GAGAACACAAAGATTGCG 102

RESULT 35
US-10-425-114-3291
; Sequence 3291, Application US/10425114
; Publication No. US20040034688A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 3291
; LENGTH: 452
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: 700258230_FLI
; US-10-425-114-3291

Query Match      82.2%; Score 14.8; DB 16; Length 452;
Best Local Similarity 88.9%; Pred. No. 1.6e+03;
```

Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GAGAACACAGATTGCG 18
|||
DB 187 GAGAACACAGATTGCG 204

RESULT 36
US-10-767-701-26974
; Sequence 26974, Application US/10767701
; Publication No. US2004017684A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement
; FILE REFERENCE: 38-21(53535)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 26974
; LENGTH: 462
; TYPE: DNA
; ORGANISM: Sorghum bicolor
; FEATURE:
; OTHER INFORMATION: Clone ID: 6672780
US-10-767-701-26974

Query Match 82.2%; Score 14.8; DB 17; Length 462;
Best Local Similarity 88.9%; Pred. No. 1.6e+03;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GAGAACACAGATTGCG 18
|||
DB 171 GAGAACACAGATTGCG 188

RESULT 37
US-09-918-995-1845
; Sequence 1845, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1845
; LENGTH: 465
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(465)
; OTHER INFORMATION: n = A,T,C or G
US-09-918-995-1845

Query Match 82.2%; Score 14.8; DB 10; Length 465;
Best Local Similarity 88.9%; Pred. No. 1.6e+03;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GAGAACACAGATTGCG 18
|||
DB 329 GAGAACACAGATTGCG 346

RESULT 38

US-10-260-238-4785/c
; Sequence 4785, Application US/10260238
; Publication No. US20040016025A1
; GENERAL INFORMATION:
; APPLICANT: Budworth, Paul R.
; APPLICANT: Moughamer, Todd G.
; APPLICANT: Briggs, Steven P.
; APPLICANT: Cooper, Bret
; APPLICANT: Glazebrook, Jane
; APPLICANT: Goff, Stephen A.
; APPLICANT: Katagiri, Fumiyaki
; APPLICANT: Krebs, Joel
; APPLICANT: Provart, Nicholas
; APPLICANT: Riche, Darrell
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: PROMOTERS FOR REGULATION OF PLANT EXPRESSION
; FILE REFERENCE: 60111-NP
; CURRENT APPLICATION NUMBER: US/10/260,238
; CURRENT FILING DATE: 2002-09-26
; PRIOR APPLICATION NUMBER: US 60/325,448
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: US 60/325,277
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: US 60/370,620
; PRIOR FILING DATE: 2002-04-04
; NUMBER OF SEQ ID NOS: 6077
; SEQ ID NO 4785
; LENGTH: 467
; TYPE: DNA
; ORGANISM: Musa acuminata
US-10-260-238-4785

Query Match 82.2%; Score 14.8; DB 16; Length 467;
Best Local Similarity 88.9%; Pred. No. 1.6e+03;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GAGAACACAGATTGCG 18
|||
DB 104 GAGAACACAGATTGCG 87

RESULT 39
US-10-425-115-81676/c
; Sequence 81676, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 81676
; LENGTH: 470
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_174503C.1
US-10-425-115-81676

Query Match 82.2%; Score 14.8; DB 18; Length 470;
Best Local Similarity 88.9%; Pred. No. 1.6e+03;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GAGAACACAGATTGCG 18
|||
DB 297 GAGAACACAGATTGCG 280

```
RESULT 40
US-10-027-632-68850/c
; Sequence 68850, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 68850
; LENGTH: 473
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-68850

Query Match
Best Local Similarity 82.2%; Score 14.8; DB 13; Length 473;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      1 GAGACACACAGATTGCC 18
Db      248 GAGATCAACAGATGCCG 231

RESULT 41
US-10-027-632-68851/c
; Sequence 68851, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 68851
; LENGTH: 473
; TYPE: DNA
; ORGANISM: Human
```

```
US-10-027-632-68851
Query Match
Best Local Similarity 82.2%; Score 14.8; DB 13; Length 473;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      1 GAGACACACAGATTGCC 18
Db      248 GAGATCAACAGATGCCG 231

RESULT 42
US-10-027-632-68850/c
; Sequence 68850, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 68850
; LENGTH: 473
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-68850

Query Match
Best Local Similarity 82.2%; Score 14.8; DB 15; Length 473;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      1 GAGACACACAGATTGCC 18
Db      248 GAGATCAACAGATGCCG 231

RESULT 43
US-10-027-632-68851/c
; Sequence 68851, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
```

```

; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 68851
; LENGTH: 473
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-68851

Query Match      82.2%; Score 14.8; DB 15; Length 473;
Best Local Similarity 88.9%; Pred. No. 1.7e+03;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GAGAACCAACAAGATTCCG 18
DB 248 GAGATCAACAAGATCGCG 231

RESULT 44
US-10-424-599-16732
; Sequence 16732, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 16732
; LENGTH: 480
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_115114C.1
US-10-424-599-16732

Query Match      82.2%; Score 14.8; DB 16; Length 480;
Best Local Similarity 88.9%; Pred. No. 1.7e+03;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GAGAACCAACAAGATTCCG 18
DB 148 GAGAACCAACAAGCTTCAC 165

RESULT 45
US-10-424-599-85425/c
; Sequence 85425, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 85425
; LENGTH: 498
; TYPE: DNA
; ORGANISM: Glycine max
```

```

; OTHER INFORMATION: Clone ID: PAT_MRT3847_48153C.1
US-10-424-599-85425

Query Match      82.2%; Score 14.8; DB 16; Length 498;
Best Local Similarity 88.9%; Pred. No. 1.7e+03;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GAGAACCAACAAGATTCCG 18
DB 470 GAGTACCAACAAGATCGCG 453

RESULT 46
US-10-027-632-217259/c
; Sequence 217259, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218

RESULT 47
US-10-027-632-217260/c
; Sequence 217260, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
```

```

; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 217260
; LENGTH: 585
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-217260
```

```
Query Match      82.2%; Score 14.8; DB 13; Length 585;
Best Local Similarity 88.9%; Pred. No. 1.7e+03;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
Qy      1 GAGAACACACAGATTCCG 18
      |||||
Db      409 GAGAACACACAGATTCCC 392
```

```
RESULT 48
US-10-027-632-217259/c
; Sequence 217259, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US/10/027,632
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 217259
; LENGTH: 585
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-217259
```

```
Query Match      82.2%; Score 14.8; DB 15; Length 585;
Best Local Similarity 88.9%; Pred. No. 1.7e+03;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
Qy      1 GAGAACACACAGATTCCG 18
      |||||
Db      409 GAGAACACACAGATTCCC 392
```

```
RESULT 49
US-10-027-632-217260/c
; Sequence 217260, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
```

```

; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 217260
; LENGTH: 585
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-217260
```

```
Query Match      82.2%; Score 14.8; DB 15; Length 585;
Best Local Similarity 88.9%; Pred. No. 1.7e+03;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
Qy      1 GAGAACACACAGATTCCG 18
      |||||
Db      409 GAGAACACACAGATTCCC 392
```

```
RESULT 50
US-10-425-114-5996
; Sequence 5996, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, Steven K.
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 5996
; LENGTH: 666
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: 700550605_FLI
US-10-425-114-5996
```

```
Query Match      82.2%; Score 14.8; DB 16; Length 666;
Best Local Similarity 88.9%; Pred. No. 1.7e+03;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
Qy      1 GAGAACACACAGATTCCG 18
      |||||
Db      401 GAGAACACACAGATTCCG 418
```

```
Search completed: December 3, 2004, 07:42:25
Job time : 1601.42 secs
```

This Page Blank (uspio)

Mon Dec 6 12:24:34 2004

us-10-050-189a-6.rn1

Page 1

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 3, 2004, 01:45:05 ; Search time 59.6842 seconds
(without alignments)
214.365 Million cell updates/sec

Title: US-10-050-189A-6

Perfect score: 18 gagacaacacagatcgc 18

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 824507 seqs, 355394441 residues

Total number of hits satisfying chosen parameters: 1649014

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 1000 summaries

Database : Issued Patents, NA: *
1: /cgn2_6/ptodata/1/ina/5A.COMB.seq: *
2: /cgn2_6/ptodata/1/ina/5B.COMB.seq: *
3: /cgn2_6/ptodata/1/ina/6A.COMB.seq: *
4: /cgn2_6/ptodata/1/ina/6B.COMB.seq: *
5: /cgn2_6/ptodata/1/ina/PTCUTS.COMB.seq: *
6: /cgn2_6/ptodata/1/ina/backfile1.seq: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	15.4	85.6	534	US-09-134-000C-2523	Sequence 2523, App
2	15.4	85.6	3178	US-08-968-563-7	Sequence 7, Appl
3	15.4	85.6	3178	US-08-969-663A-7	Sequence 7, Appl
4	15.4	85.6	3178	US-09-297-928-3	Sequence 3, Appl
5	15.4	83.3	485	US-09-276-599-4	Sequence 4, Appl
6	15.4	83.3	485	US-09-598-401C-4	Sequence 4, Appl
7	15.4	83.3	1359	US-09-248-796A-6542	Sequence 6542, Ap
8	14.8	82.2	676	US-09-495-050A-15	Sequence 15, Appl
9	14.8	82.2	1098	US-09-270-757-10453	Sequence 10453, A
10	14.8	82.2	2727	US-09-328-352-3042	Sequence 3042, Ap
11	14.8	82.2	44377	US-08-804-227C-7	Sequence 7, Appl
12	14.8	82.2	44377	US-08-804-158-1	Sequence 1, Appl
13	14.8	82.2	4403765	US-09-103-840A-2	Sequence 2, Appl
14	14.8	82.2	4411529	US-09-248-796A-13022	Sequence 13022, A
15	14.4	80.0	204	US-09-252-991A-2932	Sequence 2932, Ap
16	14.4	80.0	684	US-09-252-991A-2843	Sequence 2843, Ap
17	14.4	80.0	771	US-09-270-757-13143	Sequence 13143, A
18	14.4	80.0	783	US-09-577-934A-1	Sequence 1, Appl
19	14.4	80.0	867	US-09-252-991A-3123	Sequence 3123, Ap
20	14.4	80.0	1462	US-08-956-171E-771	Sequence 771, App
21	14.4	80.0	1462	US-08-781-986A-771	Sequence 771, App
22	14.4	80.0	1823	US-09-620-312D-516	Sequence 516, App
23	14.4	80.0	2853	US-09-556-601-25	Sequence 25, App
24	14.4	80.0	3296	US-09-907-794A-310	Sequence 310, App
25	14.4	80.0	3296	US-09-905-125A-310	Sequence 310, App
26	14.4	80.0	3296	US-09-902-775A-310	Sequence 310, App
27	14.4	80.0	3296	US-09-902-775A-310	Sequence 310, App

28	14.4	80.0	3296	US-09-906-700-310	Sequence 310, App
29	14.4	80.0	3296	US-10-140-002-369	Sequence 369, App
30	14.4	80.0	3296	US-09-903-603A-310	Sequence 310, App
31	14.4	80.0	4488	US-08-956-171E-228	Sequence 228, App
32	14.4	80.0	4488	US-08-781-986A-228	Sequence 228, App
33	14.4	80.0	162450	US-09-345-882-1	Sequence 1, Appl
34	14.4	80.0	4403765	US-09-103-840A-2	Sequence 2, Appl
35	14.4	80.0	4411529	US-09-248-796A-15780	Sequence 15780, A
36	14.4	80.0	723	US-09-252-991A-15910	Sequence 15910, A
37	14.4	77.8	723	US-09-248-796A-662	Sequence 662, App
38	14.4	77.8	835	US-09-252-991A-15910	Sequence 15910, A
39	14.4	77.8	969	US-09-248-796A-10471	Sequence 10471, A
40	14.4	77.8	1093	US-09-252-991A-15881	Sequence 15881, A
41	14.4	77.8	1317	US-09-252-991A-15836	Sequence 15836, A
42	14.4	77.8	2058	US-09-252-991A-15811	Sequence 15811, A
43	14.4	77.8	2628	US-09-609-665A-39	Sequence 39, Appl
44	13.8	76.7	66	US-08-956-171E-1643	Sequence 1643, Ap
45	13.8	76.7	66	US-08-781-986A-1643	Sequence 1643, Ap
46	13.8	76.7	160	US-09-222-575-44	Sequence 44, Appl
47	13.8	76.7	160	US-09-389-681-44	Sequence 44, Appl
48	13.8	76.7	160	US-09-320-405B-44	Sequence 44, Appl
49	13.8	76.7	160	US-09-339-338-44	Sequence 44, Appl
50	13.8	76.7	160	US-09-433-826B-44	Sequence 44, Appl
51	13.8	76.7	160	US-09-604-287A-44	Sequence 44, Appl
52	13.8	76.7	160	US-09-285-480-44	Sequence 44, Appl
53	13.8	76.7	160	US-09-834-759-44	Sequence 44, Appl
54	13.8	76.7	160	US-09-590-751A-44	Sequence 44, Appl
55	13.8	76.7	223	US-09-513-999C-16358	Sequence 16358, A
56	13.8	76.7	223	US-09-313-294A-3320	Sequence 3320, Ap
57	13.8	76.7	289	US-09-313-294A-2757	Sequence 2757, Ap
58	13.8	76.7	301	US-09-313-294A-4375	Sequence 4375, Ap
59	13.8	76.7	334	US-09-107-532A-1205	Sequence 1205, Ap
60	13.8	76.7	332	US-09-134-001C-146	Sequence 146, App
61	13.8	76.7	433	US-09-710-279-3143	Sequence 3143, Ap
62	13.8	76.7	543	US-09-252-991A-2228	Sequence 2228, Ap
63	13.8	76.7	543	US-09-270-767-14473	Sequence 14473, A
64	13.8	76.7	546	US-09-601-198-161	Sequence 161, App
65	13.8	76.7	723	US-09-270-767-1287	Sequence 1287, Ap
66	13.8	76.7	733	US-09-270-767-16569	Sequence 16569, A
67	13.8	76.7	869	US-09-071-035-359	Sequence 359, App
68	13.8	76.7	921	US-09-252-991A-2302	Sequence 2302, Ap
69	13.8	76.7	943	US-09-270-767-7830	Sequence 7830, Ap
70	13.8	76.7	943	US-09-270-767-23112	Sequence 23112, A
71	13.8	76.7	978	US-09-107-532A-2888	Sequence 2888, Ap
72	13.8	76.7	999	US-09-134-000C-2256	Sequence 2256, Ap
73	13.8	76.7	1002	US-09-071-035-357	Sequence 357, App
74	13.8	76.7	1041	US-09-252-991A-2610	Sequence 2610, App
75	13.8	76.7	1144	US-08-578-551-1	Sequence 1, Appl
76	13.8	76.7	1144	US-09-190-982-1	Sequence 1, Appl
77	13.8	76.7	1124	US-09-408-257-1	Sequence 1, Appl
78	13.8	76.7	1271	US-09-270-767-11058	Sequence 11058, A
79	13.8	76.7	1302	US-09-252-991A-12353	Sequence 12353, A
80	13.8	76.7	1359	US-09-134-001C-1136	Sequence 1136, Ap
81	13.8	76.7	1392	US-09-248-796A-4402	Sequence 4402, Ap
82	13.8	76.7	1419	US-09-634-238-68	Sequence 68, Appl
83	13.8	76.7	1589	US-09-071-035-87	Sequence 87, Appl
84	13.8	76.7	1597	US-09-270-767-12423	Sequence 12423, A
85	13.8	76.7	1611	US-09-252-991A-12304	Sequence 12304, A
86	13.8	76.7	1677	US-09-134-000C-2844	Sequence 2844, Ap
87	13.8	76.7	1677	US-09-071-035-85	Sequence 85, Appl
88	13.8	76.7	1680	US-09-252-991A-12194	Sequence 12194, A
89	13.8	76.7	1680	US-09-252-991A-11	Sequence 11, Appl
90	13.8	76.7	1695	US-09-673-395A-11	Sequence 11, Appl
91	13.8	76.7	1717	US-09-248-796A-1650	Sequence 1650, App
92	13.8	76.7	1722	US-09-620-312D-793	Sequence 793, App
93	13.8	76.7	1854	US-09-107-532A-1006	Sequence 1006, App
94	13.8	76.7	2046	US-09-252-991A-12260	Sequence 12260, A
95	13.8	76.7	2055	US-08-487-748A-3	Sequence 3, Appl
96	13.8	76.7	2055	US-08-480-070C-3	Sequence 3, Appl
97	13.8	76.7	2055	US-08-829-525-3	Sequence 3, Appl
98	13.8	76.7	2055	US-08-609-583A-3	Sequence 3, Appl
99	13.8	76.7	2055	US-08-609-583A-3	Sequence 3, Appl
100	13.8	76.7	2055	US-08-609-583A-3	Sequence 3, Appl

101	13.8	76.7	2055	4	US-09-310-367-3	Sequence 3, Appl	C 174	13.4	74.4	1806	4	US-09-134-000C-2378	Sequence 2378, Ap
102	13.8	76.7	2055	4	US-09-033-337-3	Sequence 3, Appl	175	13.4	74.4	1809	4	US-09-248-796A-6111	Sequence 6111, Ap
103	13.8	76.7	2055	4	US-09-464-231-3	Sequence 3, Appl	176	13.4	74.4	1842	4	US-09-248-796A-2421	Sequence 2421, Ap
104	13.8	76.7	2082	1	US-09-248-796A-12041	Sequence 12041, A	177	13.4	74.4	1854	4	US-09-023-655-1189	Sequence 1189, Ap
105	13.8	76.7	2781	1	US-08-178-477B-42	Sequence 42, Appl	178	13.4	74.4	1884	4	US-09-023-655-1189	Sequence 1189, Ap
106	13.8	76.7	2808	4	US-09-107-532A-99	Sequence 99, Appl	179	13.4	74.4	2025	3	US-08-992-0128-22	Sequence 23, Appl
107	13.8	76.7	3294	4	US-09-710-279-4337	Sequence 4337, Ap	180	13.4	74.4	2052	3	US-08-630-816A-45	Sequence 45, Appl
108	13.8	76.7	3285	4	US-09-710-279-535	Sequence 3535, Ap	181	13.4	74.4	2052	3	US-09-657-881A-3	Sequence 3, Appl
109	13.8	76.7	3429	1	US-08-097-972A-10	Sequence 10, Appl	182	13.4	74.4	2106	4	US-09-710-279-709	Sequence 709, App
110	13.8	76.7	3429	3	US-08-665-574C-10	Sequence 10, Appl	183	13.4	74.4	2127	3	US-09-134-001C-1178	Sequence 178, App
111	13.8	76.7	3501	4	US-10-140-002-37	Sequence 37, Appl	184	13.4	74.4	2409	1	US-09-252-991A-1551	Sequence 1551, A
112	13.8	76.7	3541	4	US-09-963-137-201	Sequence 201, App	185	13.4	74.4	2409	1	US-07-911-531-18	Sequence 18, Appl
113	13.8	76.7	3541	4	US-09-963-137-160	Sequence 160, App	186	13.4	74.4	2409	1	US-07-693-636A-18	Sequence 18, Appl
114	13.8	76.7	3541	4	US-09-023-655-1205	Sequence 1205, Ap	187	13.4	74.4	2487	3	US-08-481-435-1	Sequence 1, Appl
115	13.8	76.7	3794	4	US-09-710-279-3705	Sequence 3705, Ap	188	13.4	74.4	2487	4	US-09-543-681A-3999	Sequence 3999, Ap
116	13.8	76.7	3969	4	US-09-248-796A-5440	Sequence 5440, Ap	189	13.4	74.4	2710	4	US-08-956-171E-421	Sequence 421, App
117	13.8	76.7	4234	1	US-08-446-038B-1	Sequence 1, Appl	190	13.4	74.4	2710	4	US-08-781-986A-421	Sequence 421, App
118	13.8	76.7	4234	1	US-08-446-038B-1	Sequence 1, Appl	191	13.4	74.4	3177	4	US-09-710-279-3339	Sequence 3339, Ap
119	13.8	76.7	4234	1	US-08-805-445-1	Sequence 1, Appl	192	13.4	74.4	3828	4	US-09-710-279-3391	Sequence 3391, Ap
120	13.8	76.7	4234	2	US-08-064-067D-1	Sequence 1, Appl	193	13.4	74.4	4302	4	US-09-543-681A-3170	Sequence 3170, Ap
121	13.8	76.7	4234	2	US-09-066-208-1	Sequence 1, Appl	194	13.4	74.4	4866	4	US-08-956-171E-276	Sequence 276, App
122	13.8	76.7	4234	2	US-09-331-581-1	Sequence 1, Appl	195	13.4	74.4	4866	4	US-08-956-171E-276	Sequence 276, App
123	13.8	76.7	6256	2	US-08-475-891A-1	Sequence 1, Appl	196	13.4	74.4	4866	4	US-08-956-171E-276	Sequence 276, App
124	13.8	76.7	6256	2	US-08-475-891A-1	Sequence 1, Appl	197	13.4	74.4	4866	4	US-08-956-171E-276	Sequence 276, App
125	13.8	76.7	6256	2	US-08-567-375-1	Sequence 1, Appl	198	13.4	74.4	4866	4	US-08-956-171E-276	Sequence 276, App
126	13.8	76.7	6256	2	US-08-567-375-1	Sequence 1, Appl	199	13.4	74.4	4866	4	US-08-956-171E-276	Sequence 276, App
127	13.8	76.7	7286	3	US-09-331-581-3	Sequence 3, Appl	200	13.4	74.4	5537	1	US-08-483-852A-9	Sequence 9, Appl
128	13.8	76.7	7938	3	US-09-331-581-14	Sequence 14, Appl	201	13.4	74.4	5537	1	US-08-483-852A-9	Sequence 9, Appl
129	13.8	76.7	11721	3	US-09-026-039-3	Sequence 3, Appl	202	13.4	74.4	5537	2	US-08-483-852A-9	Sequence 9, Appl
130	13.8	76.7	11721	3	US-09-026-039-3	Sequence 3, Appl	203	13.4	74.4	5537	2	US-08-483-852A-9	Sequence 9, Appl
131	13.8	76.7	15397	4	US-08-956-171E-205	Sequence 205, App	204	13.4	74.4	5537	2	US-08-483-852A-9	Sequence 9, Appl
132	13.8	76.7	15397	4	US-08-956-171E-205	Sequence 205, App	205	13.4	74.4	5537	2	US-08-483-852A-9	Sequence 9, Appl
133	13.8	76.7	15397	4	US-08-956-171E-205	Sequence 205, App	206	13.4	74.4	5537	2	US-08-483-852A-9	Sequence 9, Appl
134	13.8	76.7	15397	4	US-08-956-171E-205	Sequence 205, App	207	13.4	74.4	5537	2	US-08-483-852A-9	Sequence 9, Appl
135	13.8	76.7	15397	4	US-08-956-171E-205	Sequence 205, App	208	13.4	74.4	5537	2	US-08-483-852A-9	Sequence 9, Appl
136	13.8	76.7	15397	4	US-08-956-171E-205	Sequence 205, App	209	13.4	74.4	5537	2	US-08-483-852A-9	Sequence 9, Appl
137	13.8	76.7	15397	4	US-08-956-171E-205	Sequence 205, App	210	13.4	74.4	5537	2	US-08-483-852A-9	Sequence 9, Appl
138	13.8	76.7	15397	4	US-08-956-171E-205	Sequence 205, App	211	13.4	74.4	5537	2	US-08-483-852A-9	Sequence 9, Appl
139	13.8	76.7	15397	4	US-08-956-171E-205	Sequence 205, App	212	13.4	74.4	5537	2	US-08-483-852A-9	Sequence 9, Appl
140	13.8	76.7	15397	4	US-08-956-171E-205	Sequence 205, App	213	13.4	74.4	5537	2	US-08-483-852A-9	Sequence 9, Appl
141	13.8	76.7	15397	4	US-08-956-171E-205	Sequence 205, App	214	13.4	74.4	5537	2	US-08-483-852A-9	Sequence 9, Appl
142	13.8	76.7	15397	4	US-08-956-171E-205	Sequence 205, App	215	13.4	74.4	5537	2	US-08-483-852A-9	Sequence 9, Appl
143	13.8	76.7	15397	4	US-08-956-171E-205	Sequence 205, App	216	13.4	74.4	5537	2	US-08-483-852A-9	Sequence 9, Appl
144	13.8	76.7	15397	4	US-08-956-171E-205	Sequence 205, App	217	13.4	74.4	5537	2	US-08-483-852A-9	Sequence 9, Appl
145	13.8	76.7	15397	4	US-08-956-171E-205	Sequence 205, App	218	13.4	74.4	5537	2	US-08-483-852A-9	Sequence 9, Appl
146	13.8	76.7	15397	4	US-08-956-171E-205	Sequence 205, App	219	13.4	74.4	5537	2	US-08-483-852A-9	Sequence 9, Appl
147	13.8	76.7	15397	4	US-08-956-171E-205	Sequence 205, App	220	13.4	74.4	5537	2	US-08-483-852A-9	Sequence 9, Appl
148	13.8	76.7	15397	4	US-08-956-171E-205	Sequence 205, App	221	13.4	74.4	5537	2	US-08-483-852A-9	Sequence 9, Appl
149	13.8	76.7	15397	4	US-08-956-171E-205	Sequence 205, App	222	13.4	74.4	5537	2	US-08-483-852A-9	Sequence 9, Appl
150	13.8	76.7	15397	4	US-08-956-171E-205	Sequence 205, App	223	13.4	74.4	5537	2	US-08-483-852A-9	Sequence 9, Appl
151	13.8	76.7	15397	4	US-08-956-171E-205	Sequence 205, App	224	13.4	74.4	5537	2	US-08-483-852A-9	Sequence 9, Appl
152	13.8	76.7	15397	4	US-08-956-171E-205	Sequence 205, App	225	13.4	74.4	5537	2	US-08-483-852A-9	Sequence 9, Appl
153	13.8	76.7	15397	4	US-08-956-171E-205	Sequence 205, App	226	13.4	74.4	5537	2	US-08-483-852A-9	Sequence 9, Appl
154	13.8	76.7	15397	4	US-08-956-171E-205	Sequence 205, App	227	13.4	74.4	5537	2	US-08-483-852A-9	Sequence 9, Appl
155	13.8	76.7	15397	4	US-08-956-171E-205	Sequence 205, App	228	13.4	74.4	5537	2	US-08-483-852A-9	Sequence 9, Appl
156	13.8	76.7	15397	4	US-08-956-171E-205	Sequence 205, App	229	13.4	74.4	5537	2	US-08-483-852A-9	Sequence 9, Appl
157	13.8	76.7	15397	4	US-08-956-171E-205	Sequence 205, App	230	13.4	74.4	5537	2	US-08-483-852A-9	Sequence 9, Appl
158	13.8	76.7	15397	4	US-08-956-171E-205	Sequence 205, App	231	13.4	74.4	5537	2	US-08-483-852A-9	Sequence 9, Appl
159	13.8	76.7	15397	4	US-08-956-171E-205	Sequence 205, App	232	13.4	74.4	5537	2	US-08-483-852A-9	Sequence 9, Appl
160	13.8	76.7	15397	4	US-08-956-171E-205	Sequence 205, App	233	13.4	74.4	5537	2	US-08-483-852A-9	Sequence 9, Appl
161	13.8	76.7	15397	4	US-08-956-171E-205	Sequence 205, App	234	13.4	74.4	5537	2	US-08-483-852A-9	Sequence 9, Appl
162	13.8	76.7	15397	4	US-08-956-171E-205	Sequence 205, App	235	13.4	74.4	5537	2	US-08-483-852A-9	Sequence 9, Appl
163	13.8	76.7	15397	4	US-08-956-171E-205	Sequence 205, App	236	13.4	74.4	5537	2	US-08-483-852A-9	Sequence 9, Appl
164	13.8	76.7	15397	4	US-08-956-171E-205	Sequence 205, App	237	13.4	74.4	5537	2	US-08-483-852A-9	Sequence 9, Appl
165	13.8	76.7	15397	4	US-08-956-171E-205	Sequence 205, App	238	13.4	74.4	5537	2	US-08-483-852A-9	Sequence 9, Appl
166	13.8	76.7	15397	4	US-08-956-171E-205	Sequence 205, App	239	13.4	74.4	5537	2	US-08-483-852A-9	Sequence 9, Appl
167	13.8	76.7	15397	4	US-08-956-171E-205	Sequence 205, App	240	13.4	74.4	5537	2	US-08-483-852A-9	Sequence 9, Appl
168	13.8	76.7	15397	4	US-08-956-171E-205	Sequence 205, App	241	13.4	74.4	5537	2	US-08-483-852A-9	Sequence 9, Appl
169	13.8	76.7	15397	4	US-08-956-171E-205	Sequence 205, App	242	13.4	74.4	5537	2	US-08-483-852A-9	Sequence 9, Appl
170	13.8	76.7	15397	4	US-08-956-171E-205	Sequence 205, App	243	13.4	74.4	5537	2	US-08-483-852A-9	Sequence 9, Appl
171	13.8	76.7	15397	4	US-08-956-171E-205	Sequence 205, App	244	13.4	74.4	5537	2	US-08-483-852A-9	Sequence 9, Appl
172	13.8	76.7	15397	4	US-08-956-171E-205	Sequence 205, App	245	13.4	74.4	5537	2	US-08-483-852A-9	Sequence 9, Appl
173	13.8	76.7	15397	4	US-08-956-171E-205	Sequence 205, App	246	13.4	74.4	5537	2	US-08-483-852A-9	Sequence 9, Appl

C 247	13.2	73.3	620	4	US-09-270-767-26488	Sequence 28488, A	320	13.2	73.3	2521	1	US-08-332-638-59	Sequence 59, Appl
C 248	13.2	73.3	624	1	US-09-540-236-145	Sequence 145, App	321	13.2	73.3	2525	3	US-08-294-312B-1	Sequence 1, Appl
C 249	13.2	73.3	662	1	US-07-807-043B-10	Sequence 10, Appl	322	13.2	73.3	2525	4	US-08-468-024B-1	Sequence 1, Appl
C 250	13.2	73.3	662	2	US-08-299-849B-10	Sequence 10, Appl	323	13.2	73.3	2525	4	US-08-187-757D-1	Sequence 1, Appl
C 251	13.2	73.3	662	1	US-08-142-168A-10	Sequence 10, Appl	324	13.2	73.3	2525	4	US-08-465-679-1	Sequence 1, Appl
C 252	13.2	73.3	662	3	US-08-967-727-10	Sequence 10, Appl	325	13.2	73.3	2525	4	US-08-210-143C-1	Sequence 1, Appl
C 253	13.2	73.3	662	3	US-08-037-230D-10	Sequence 10, Appl	326	13.2	73.3	2580	2	US-08-887-798-1	Sequence 1, Appl
C 254	13.2	73.3	662	4	US-09-583-850-10	Sequence 10, Appl	327	13.2	73.3	2908	4	US-09-930-181-1	Sequence 1, Appl
C 255	13.2	73.3	662	4	US-09-579-197-10	Sequence 10, Appl	328	13.2	73.3	2943	1	US-08-042-747A-7	Sequence 7, Appl
C 256	13.2	73.3	662	4	US-09-404-026-10	Sequence 10, Appl	329	13.2	73.3	2980	4	US-10-003-690-1	Sequence 1, Appl
C 257	13.2	73.3	662	4	US-09-312-164-10	Sequence 10, Appl	330	13.2	73.3	2982	4	US-09-710-279-385	Sequence 385, App
C 258	13.2	73.3	662	4	US-09-583-110-960	Sequence 960, App	331	13.2	73.3	3045	4	US-09-633-238-10	Sequence 10, Appl
C 259	13.2	73.3	703	3	US-08-998-416-632	Sequence 632, App	332	13.2	73.3	3120	4	US-09-432-468A-14	Sequence 14, Appl
C 260	13.2	73.3	705	4	US-09-134-000C-2631	Sequence 2631, App	333	13.2	73.3	3177	1	US-08-042-747A-4	Sequence 4, Appl
C 261	13.2	73.3	815	4	US-08-482-709A-5	Sequence 5, Appl	334	13.2	73.3	3277	4	US-09-293-549-5	Sequence 3, Appl
C 262	13.2	73.3	815	4	US-09-248-766A-2554	Sequence 2554, App	335	13.2	73.3	3364	4	US-09-930-181-3	Sequence 3, Appl
C 263	13.2	73.3	898	4	US-09-270-767-13110	Sequence 13110, A	336	13.2	73.3	3472	6	US-09-620-665A-84	Sequence 84, Appl
C 264	13.2	73.3	909	1	US-08-701-952A-3	Sequence 3, Appl	337	13.2	73.3	3494	4	US-08-956-171E-304	Sequence 304, App
C 265	13.2	73.3	957	4	US-09-543-681A-2212	Sequence 2212, App	338	13.2	73.3	3821	4	US-08-781-986A-304	Sequence 880, App
C 266	13.2	73.3	957	4	US-09-489-039A-3737	Sequence 3737, App	339	13.2	73.3	4017	4	US-10-101-464A-880	Sequence 830, App
C 267	13.2	73.3	993	4	US-09-270-767-7326	Sequence 7326, App	340	13.2	73.3	4056	4	US-09-710-279-3879	Sequence 3879, App
C 268	13.2	73.3	1000	4	US-09-270-767-7326	Sequence 22608, A	341	13.2	73.3	4158	4	US-09-252-991A-12136	Sequence 12126, A
C 269	13.2	73.3	1000	4	US-09-270-767-7326	Sequence 5, Appl	342	13.2	73.3	4884	4	US-09-252-991A-12292	Sequence 12292, A
C 270	13.2	73.3	1017	3	US-09-310-611-5	Sequence 12982, A	343	13.2	73.3	4884	4	US-09-073-354-6	Sequence 6, Appl
C 271	13.2	73.3	1050	4	US-09-252-991A-12982	Sequence 13142, A	344	13.2	73.3	5339	3	US-08-656-005A-5	Sequence 6, Appl
C 272	13.2	73.3	1092	4	US-09-252-991A-13142	Sequence 13142, A	345	13.2	73.3	5339	3	US-09-073-259-6	Sequence 6, Appl
C 273	13.2	73.3	1146	4	US-09-710-279-3055	Sequence 3055, App	346	13.2	73.3	5339	3	US-09-073-259-6	Sequence 6, Appl
C 274	13.2	73.3	1182	4	US-09-252-991A-12239	Sequence 12239, App	347	13.2	73.3	5339	3	US-09-073-259-6	Sequence 6, Appl
C 275	13.2	73.3	1197	4	US-08-489-039A-6585	Sequence 6585, App	348	13.2	73.3	5339	3	US-09-363-095-6	Sequence 6, Appl
C 276	13.2	73.3	1344	4	US-09-252-991A-6450	Sequence 6450, App	349	13.2	73.3	5342	3	US-09-418-027-6	Sequence 1, Appl
C 277	13.2	73.3	1366	4	US-09-270-767-12957	Sequence 12957, A	350	13.2	73.3	5342	3	US-08-902-633-1	Sequence 1, Appl
C 278	13.2	73.3	1440	1	US-07-936-163-1	Sequence 1, Appl	351	13.2	73.3	5342	3	US-09-073-354-5	Sequence 5, Appl
C 279	13.2	73.3	1506	4	US-09-489-039A-6807	Sequence 6807, App	352	13.2	73.3	5342	3	US-08-656-005A-5	Sequence 5, Appl
C 280	13.2	73.3	1506	4	US-09-252-991A-6098	Sequence 6098, App	353	13.2	73.3	5342	3	US-09-073-259-5	Sequence 5, Appl
C 281	13.2	73.3	1560	3	US-09-105-039A-1	Sequence 1, Appl	354	13.2	73.3	5342	3	US-09-363-095-5	Sequence 5, Appl
C 282	13.2	73.3	1560	3	US-09-620-312D-974	Sequence 974, App	355	13.2	73.3	5342	3	US-09-418-027-5	Sequence 5, Appl
C 283	13.2	73.3	1586	4	US-09-352-772-3	Sequence 3, Appl	356	13.2	73.3	5342	3	US-08-404-445-1	Sequence 1, Appl
C 284	13.2	73.3	1589	4	US-09-270-767-13252	Sequence 13252, A	357	13.2	73.3	7060	4	US-09-221-011B-479	Sequence 479, App
C 285	13.2	73.3	1593	4	US-09-134-000C-1665	Sequence 1665, App	358	13.2	73.3	7792	4	US-09-220-123-24	Sequence 12, Appl
C 286	13.2	73.3	1599	4	US-09-252-991A-11471	Sequence 11471, A	359	13.2	73.3	7796	4	US-09-521-511C-12	Sequence 12, Appl
C 287	13.2	73.3	1605	3	US-09-105-039A-3	Sequence 3, Appl	360	13.2	73.3	13365	3	US-08-961-527-111	Sequence 17, App
C 288	13.2	73.3	1607	4	US-09-091-097-9	Sequence 9, Appl	361	13.2	73.3	13365	3	US-09-453-702B-18	Sequence 70, Appl
C 289	13.2	73.3	1674	4	US-09-543-681A-2434	Sequence 2434, App	362	13.2	73.3	17710	4	US-09-956-004-70	Sequence 18, Appl
C 290	13.2	73.3	1740	3	US-09-252-991A-12801	Sequence 12801, A	363	13.2	73.3	33000	3	US-09-215-694-18	Sequence 10, Appl
C 291	13.2	73.3	1752	4	US-09-188-910-225	Sequence 225, App	364	13.2	73.3	33000	3	US-09-567-668-10	Sequence 31, Appl
C 292	13.2	73.3	1752	4	US-09-312-283C-225	Sequence 225, App	365	13.2	73.3	33000	3	US-09-567-668-10	Sequence 32, Appl
C 293	13.2	73.3	1770	4	US-09-252-991A-12265	Sequence 12265, App	366	13.2	73.3	33000	3	US-09-497-855A-32	Sequence 11, Appl
C 294	13.2	73.3	1797	4	US-09-252-991A-12418	Sequence 12418, A	367	13.2	73.3	33000	4	US-10-172-911-11	Sequence 20, Appl
C 295	13.2	73.3	1830	4	US-09-774-528-221	Sequence 221, App	368	13.2	73.3	246240	2	US-08-724-394A-21	Sequence 21, Appl
C 296	13.2	73.3	1830	4	US-09-557-921-1	Sequence 1, Appl	369	13.2	73.3	246240	2	US-08-724-394A-22	Sequence 22, Appl
C 297	13.2	73.3	1830	4	US-09-630-983A-2	Sequence 2, Appl	370	13.2	73.3	246240	2	US-08-916-421B-1	Sequence 1, Appl
C 298	13.2	73.3	2000	4	US-09-516-052-23	Sequence 23, Appl	371	13.2	73.3	246240	2	US-09-692-570-1	Sequence 1, Appl
C 299	13.2	73.3	2015	4	US-09-774-528-198	Sequence 358, App	372	13.2	73.3	1664976	4	US-09-134-001C-2386	Sequence 2386, App
C 300	13.2	73.3	2028	4	US-09-270-767-12056	Sequence 12056, App	373	13.2	73.3	284	3	US-09-513-294A-1202	Sequence 55, Appl
C 301	13.2	73.3	2077	1	US-08-217-337-7	Sequence 7, Appl	374	13.2	73.3	284	3	US-09-513-294A-1202	Sequence 24, Appl
C 302	13.2	73.3	2112	4	US-10-116-336-5	Sequence 5, Appl	375	13.2	73.3	284	3	US-09-513-294A-1202	Sequence 24, Appl
C 303	13.2	73.3	2214	4	US-09-248-766A-216	Sequence 216, App	376	13.2	73.3	322	3	US-09-851-994-55	Sequence 24, Appl
C 304	13.2	73.3	2249	4	US-09-270-767-12680	Sequence 12680, A	377	13.2	73.3	322	3	US-09-851-994-55	Sequence 24, Appl
C 305	13.2	73.3	2289	4	US-10-116-336-3	Sequence 3, Appl	378	13.2	73.3	322	3	US-09-851-994-55	Sequence 24, Appl
C 306	13.2	73.3	2334	4	US-09-602-787A-181	Sequence 181, App	379	13.2	73.3	322	3	US-09-851-994-55	Sequence 24, Appl
C 307	13.2	73.3	2334	4	US-10-003-690-3	Sequence 3, Appl	380	13.2	73.3	322	3	US-09-851-994-55	Sequence 24, Appl
C 308	13.2	73.3	2337	4	US-10-116-336-1	Sequence 1, Appl	381	13.2	73.3	322	3	US-09-851-994-55	Sequence 24, Appl
C 309	13.2	73.3	2364	4	US-09-252-991A-12368	Sequence 12368, A	382	13.2	73.3	322	3	US-09-851-994-55	Sequence 24, Appl
C 310	13.2	73.3	2385	4	US-08-352-902D-145	Sequence 145, App	383	13.2	73.3	322	3	US-09-851-994-55	Sequence 24, Appl
C 311	13.2	73.3	2484	2	US-08-209-521-8	Sequence 8, Appl	384	13.2	73.3	322	3	US-09-851-994-55	Sequence 24, Appl
C 312	13.2	73.3	2484	3	US-08-961-810-4	Sequence 4, Appl	385	13.2	73.3	322	3	US-09-851-994-55	Sequence 24, Appl
C 313	13.2	73.3	2484	3	US-08-352-903B-4	Sequence 4, Appl	386	13.2	73.3	322	3	US-09-851-994-55	Sequence 24, Appl
C 314	13.2	73.3	2484	4	US-09-265-503B-4	Sequence 4, Appl	387	13.2	73.3	322	3	US-09-851-994-55	Sequence 24, Appl
C 315	13.2	73.3	2484	4	US-09-708-200-16	Sequence 16, Appl	388	13.2	73.3	322	3	US-09-851-994-55	Sequence 24, Appl
C 316	13.2	73.3	2484	4	US-09-788-657-10	Sequence 10, Appl	389	13.2	73.3	322	3	US-09-851-994-55	Sequence 24, Appl
C 317	13.2	73.3	2484	4	US-09-712-691-14	Sequence 14, Appl	390	13.2	73.3	322	3	US-09-851-994-55	Sequence 24, Appl
C 318	13.2	73.3	2521	1	US-08-188-228-59	Sequence 59, Appl	391	13.2	73.3	322	3	US-09-851-994-55	Sequence 24, Appl
C 319	13.2	73.3	2521	1	US-08-332-643-53	Sequence 53, Appl	392	13.2	73.3	322	3	US-09-851-994-55	Sequence 24, Appl

C 333	13	72.2	550	3	US-09-127-480-110	Sequence 110, App	466	13	72.2	2509	3	US-08-229-050-14	Sequence 14, Appl
C 334	13	72.2	550	3	US-08-496-841C-110	Sequence 110, App	467	13	72.2	2509	3	US-08-801-563-14	Sequence 14, Appl
C 335	13	72.2	550	3	US-09-124-523-110	Sequence 110, App	468	13	72.2	2509	3	US-09-710-279-4170	Sequence 4170, Ap
C 336	13	72.2	550	4	US-09-636-796A-110	Sequence 110, App	C 469	13	72.2	3049	4	US-09-710-279-4151	Sequence 4151, Ap
C 337	13	72.2	550	4	US-08-431-048F-110	Sequence 110, App	470	13	72.2	3073	4	US-09-710-279-3606	Sequence 3606, Ap
338	13	72.2	611	4	US-09-376-113-4	Sequence 4, Appl	471	13	72.2	3132	4	US-09-710-279-3455	Sequence 3455, Ap
339	13	72.2	631	4	US-09-376-113-6	Sequence 6, Appl	472	13	72.2	3371	4	US-09-710-279-3490	Sequence 3490, Ap
400	13	72.2	642	3	US-09-449-218D-9	Sequence 9, Appl	473	13	72.2	3465	4	US-09-710-279-1779	Sequence 1779, Ap
401	13	72.2	642	3	US-09-449-218D-45	Sequence 45, Appl	C 474	13	72.2	3543	4	US-09-252-991A-15893	Sequence 1593, A
402	13	72.2	642	4	US-09-668-529A-9	Sequence 9, Appl	475	13	72.2	3722	4	US-09-919-039-98	Sequence 98, Appl
403	13	72.2	642	4	US-09-668-529A-45	Sequence 45, Appl	476	13	72.2	3744	2	US-08-348-353-16	Sequence 16, Appl
404	13	72.2	642	4	US-09-668-037A-9	Sequence 9, Appl	477	13	72.2	3744	2	US-08-465-965-16	Sequence 16, Appl
405	13	72.2	642	4	US-09-668-037A-45	Sequence 45, Appl	478	13	72.2	3744	2	US-08-465-966-16	Sequence 16, Appl
406	13	72.2	679	3	US-08-821-994-54	Sequence 54, Appl	C 479	13	72.2	4103	4	US-09-710-279-3532	Sequence 3532, Ap
407	13	72.2	690	3	US-08-821-994-53	Sequence 53, Appl	480	13	72.2	4257	4	US-09-710-279-3532	Sequence 3532, Ap
C 408	13	72.2	729	3	US-08-858-207A-198	Sequence 198, App	481	13	72.2	4483	1	US-08-181-271A-5	Sequence 1596, A
C 409	13	72.2	855	4	US-09-248-796A-6820	Sequence 6820, Ap	482	13	72.2	4483	1	US-08-449-315-5	Sequence 5, Appl
C 410	13	72.2	879	4	US-09-583-110-132	Sequence 132, App	483	13	72.2	4483	1	US-08-444-803-5	Sequence 5, Appl
C 411	13	72.2	888	4	US-09-710-279-2945	Sequence 2945, Ap	484	13	72.2	4483	1	US-08-449-043-3	Sequence 5, Appl
C 412	13	72.2	1102	3	US-08-821-994-86	Sequence 86, Appl	485	13	72.2	4483	1	US-08-456-265A-5	Sequence 5, Appl
C 413	13	72.2	1125	4	US-09-328-352-336	Sequence 336, Ap	486	13	72.2	4483	1	US-08-455-416-5	Sequence 5, Appl
414	13	72.2	1197	4	US-09-614-221A-549	Sequence 549, App	487	13	72.2	4483	1	US-08-455-416-5	Sequence 5, Appl
415	13	72.2	1256	4	US-09-556-877-21	Sequence 21, Appl	488	13	72.2	4483	1	US-08-454-876-5	Sequence 5, Appl
416	13	72.2	1256	4	US-09-288-594A-21	Sequence 21, Appl	489	13	72.2	4483	1	US-08-457-364-5	Sequence 5, Appl
417	13	72.2	1256	4	US-09-620-412C-21	Sequence 21, Appl	490	13	72.2	4483	2	US-08-456-262-5	Sequence 5, Appl
418	13	72.2	1256	4	US-09-410-568-21	Sequence 21, Appl	491	13	72.2	4483	2	US-08-456-262-5	Sequence 5, Appl
419	13	72.2	1256	4	US-09-558-419-21	Sequence 21, Appl	492	13	72.2	4483	2	US-08-455-746-5	Sequence 5, Appl
420	13	72.2	1358	1	US-08-181-771A-20	Sequence 20, Appl	493	13	72.2	4483	2	US-08-455-746-5	Sequence 5, Appl
421	13	72.2	1358	1	US-08-449-315-20	Sequence 20, Appl	494	13	72.2	4483	3	US-09-350-600-5	Sequence 5, Appl
422	13	72.2	1358	1	US-08-444-803-20	Sequence 20, Appl	495	13	72.2	4483	4	US-09-906-234-5	Sequence 5, Appl
423	13	72.2	1358	1	US-08-449-043-20	Sequence 20, Appl	496	13	72.2	4699	1	US-08-181-271A-6	Sequence 6, Appl
424	13	72.2	1358	1	US-08-456-265A-20	Sequence 20, Appl	497	13	72.2	4699	1	US-08-449-315-6	Sequence 6, Appl
425	13	72.2	1358	1	US-08-455-416-20	Sequence 20, Appl	498	13	72.2	4699	1	US-08-444-803-6	Sequence 6, Appl
426	13	72.2	1358	1	US-08-455-444-20	Sequence 20, Appl	499	13	72.2	4699	1	US-08-449-043-6	Sequence 6, Appl
427	13	72.2	1358	1	US-08-454-876-20	Sequence 20, Appl	500	13	72.2	4699	1	US-08-456-265A-6	Sequence 6, Appl
428	13	72.2	1358	2	US-08-457-364-20	Sequence 20, Appl	501	13	72.2	4699	1	US-08-455-416-6	Sequence 6, Appl
429	13	72.2	1358	2	US-08-456-262-20	Sequence 20, Appl	502	13	72.2	4699	1	US-08-454-876-6	Sequence 6, Appl
430	13	72.2	1358	2	US-08-455-736-20	Sequence 20, Appl	503	13	72.2	4699	2	US-08-457-364-6	Sequence 6, Appl
431	13	72.2	1358	2	US-08-971-217-20	Sequence 20, Appl	504	13	72.2	4699	2	US-08-456-262-6	Sequence 6, Appl
432	13	72.2	1358	2	US-08-971-217-20	Sequence 20, Appl	505	13	72.2	4699	2	US-08-456-240-6	Sequence 6, Appl
433	13	72.2	1358	3	US-09-350-600-20	Sequence 20, Appl	506	13	72.2	4699	2	US-08-456-240-6	Sequence 6, Appl
434	13	72.2	1358	4	US-09-506-234-20	Sequence 20, Appl	507	13	72.2	4699	2	US-08-455-736-6	Sequence 6, Appl
C 435	13	72.2	1365	4	US-09-252-991A-16531	Sequence 16531, A	508	13	72.2	4699	2	US-08-971-217-6	Sequence 6, Appl
436	13	72.2	1390	3	US-08-821-994-61	Sequence 61, Appl	509	13	72.2	4699	3	US-09-350-600-6	Sequence 6, Appl
437	13	72.2	1434	3	US-08-821-994-62	Sequence 62, Appl	510	13	72.2	4699	4	US-09-906-234-6	Sequence 6, Appl
438	13	72.2	1441	3	US-08-821-994-63	Sequence 63, Appl	C 511	13	72.2	5304	4	US-09-023-655-664	Sequence 664, App
439	13	72.2	1474	3	US-08-821-994-64	Sequence 64, Appl	512	13	72.2	5406	4	US-09-386-962C-1	Sequence 1, Appl
C 440	13	72.2	1507	1	US-07-506-871-13	Sequence 13, Appl	513	13	72.2	5406	4	US-09-386-959-1	Sequence 1, Appl
C 441	13	72.2	1566	4	US-09-107-532A-307	Sequence 307, App	514	13	72.2	6057	4	US-09-231-899-79	Sequence 79, Appl
442	13	72.2	1632	3	US-09-134-001C-1586	Sequence 1586, Ap	C 515	13	72.2	6654	4	US-08-961-527-98	Sequence 98, Appl
443	13	72.2	1661	3	US-08-821-994-82	Sequence 82, Appl	516	13	72.2	9301	4	US-09-449-218B-18	Sequence 18, Appl
444	13	72.2	1689	3	US-09-108-010B-14	Sequence 14, Appl	517	13	72.2	9301	4	US-09-668-529A-18	Sequence 18, Appl
445	13	72.2	1689	3	US-09-758-652-14	Sequence 14, Appl	518	13	72.2	9301	4	US-09-668-037A-18	Sequence 18, Appl
446	13	72.2	1721	1	US-08-281-916-1	Sequence 1, Appl	519	13	72.2	9632	4	US-09-403-752A-1	Sequence 1, Appl
447	13	72.2	1721	2	US-08-460-725-1	Sequence 2, Appl	520	13	72.2	9639	3	US-09-147-208-26	Sequence 26, Appl
C 448	13	72.2	1905	4	US-09-552-991A-16307	Sequence 16307, A	521	13	72.2	9639	3	US-09-147-208-26	Sequence 26, Appl
C 449	13	72.2	2001	4	US-09-252-991A-15971	Sequence 15971, A	522	13	72.2	9639	3	US-09-550-117A-26	Sequence 26, Appl
C 450	13	72.2	2136	4	US-09-214-909-17	Sequence 17, Appl	523	13	72.2	10241	1	US-09-182-145-55	Sequence 55, Appl
451	13	72.2	2163	4	US-09-252-991A-15826	Sequence 15826, A	524	13	72.2	10379	1	US-08-508-448C-16	Sequence 16, Appl
452	13	72.2	2157	4	US-09-252-991A-16172	Sequence 16172, A	525	13	72.2	11284	3	US-09-182-145-54	Sequence 54, Appl
453	13	72.2	2301	3	US-09-449-218B-1	Sequence 1, Appl	526	13	72.2	11562	3	US-08-978-741-5	Sequence 5, Appl
454	13	72.2	2301	3	US-09-449-218B-3	Sequence 3, Appl	527	13	72.2	11562	3	US-08-927-317-7	Sequence 7, Appl
455	13	72.2	2301	3	US-09-449-218D-5	Sequence 5, Appl	528	13	72.2	19227	3	US-09-073-492-1	Sequence 1, Appl
456	13	72.2	2301	3	US-09-449-218D-7	Sequence 7, Appl	529	13	72.2	19227	3	US-09-090-793-13	Sequence 13, Appl
457	13	72.2	2301	4	US-09-668-529A-1	Sequence 1, Appl	530	13	72.2	28001	4	US-09-231-899-13	Sequence 13, Appl
458	13	72.2	2301	4	US-09-668-529A-3	Sequence 3, Appl	531	13	72.2	28001	4	US-09-819-993-3	Sequence 3, Appl
459	13	72.2	2301	4	US-09-668-529A-5	Sequence 5, Appl	532	13	72.2	28001	4	US-10-193-295-3	Sequence 3, Appl
460	13	72.2	2301	4	US-09-668-529A-7	Sequence 7, Appl	533	13	72.2	40138	4	US-09-090-793-12	Sequence 12, Appl
461	13	72.2	2301	4	US-09-668-037A-1	Sequence 1, Appl	534	13	72.2	40138	4	US-09-231-899-12	Sequence 12, Appl
462	13	72.2	2301	4	US-09-668-037A-3	Sequence 3, Appl	C 535	13	72.2	640681	4	US-09-437-855A-38	Sequence 38, Appl
463	13	72.2	2301	4	US-09-668-037A-5	Sequence 5, Appl	536	13	71.1	26	2	US-09-790-988-1	Sequence 1, Appl
464	13	72.2	2301	4	US-09-668-037A-7	Sequence 7, Appl	C 537	12.8	71.1	47	4	US-09-197-008-3	Sequence 3, Appl
465	13	72.2	2509	1	US-08-047-413-14	Sequence 14, Appl	538	12.8	71.1	93	4	US-09-422-978-1257	Sequence 1257, Ap
												US-09-513-999C-16427	Sequence 16427, A

539	12.8	71.1	144	4	US-09-313-294A-2557	Sequence 2557, Ap	c 612	12.8	71.1	631	4	US-09-270-767-10760	Sequence 10760, A
540	12.8	71.1	148	4	US-09-248-796A-10752	Sequence 10752, A	c 613	12.8	71.1	631	4	US-09-270-767-15278	Sequence 15278, A
541	12.8	71.1	200	4	US-09-313-294A-3705	Sequence 3705, Ap	c 614	12.8	71.1	654	4	US-09-583-110-150	Sequence 150, App
542	12.8	71.1	240	4	US-09-513-999C-33512	Sequence 33512, A	c 615	12.8	71.1	654	4	US-09-583-110-751	Sequence 751, App
543	12.8	71.1	244	5	US-08-238-821B-47	Sequence 47, Appl	c 616	12.8	71.1	654	4	US-09-583-110-1831	Sequence 1831, Ap
544	12.8	71.1	244	5	PCT-US95-05744-47	Sequence 47, Appl	c 617	12.8	71.1	670	4	US-09-270-767-10084	Sequence 10084, A
545	12.8	71.1	248	4	US-09-016-434-524	Sequence 524, App	c 618	12.8	71.1	690	2	US-08-612-840A-7	Sequence 7, Appl1
546	12.8	71.1	250	4	US-09-270-767-8208	Sequence 8208, Ap	c 619	12.8	71.1	690	2	US-09-969-763-2	Sequence 2, Appl1
547	12.8	71.1	250	4	US-09-270-767-23490	Sequence 23490, A	c 620	12.8	71.1	707	2	US-08-351-176-1	Sequence 1, Appl1
548	12.8	71.1	258	4	US-09-248-796A-11874	Sequence 11874, A	c 621	12.8	71.1	707	2	US-09-455-976-1	Sequence 1, Appl1
549	12.8	71.1	274	4	US-09-270-767-1077	Sequence 1077, Ap	c 622	12.8	71.1	707	2	US-09-222-227-1	Sequence 1, Appl1
550	12.8	71.1	274	4	US-09-270-767-16359	Sequence 16359, A	c 623	12.8	71.1	712	3	US-08-998-416-887	Sequence 887, App
551	12.8	71.1	274	4	US-09-513-999C-23822	Sequence 23822, A	c 624	12.8	71.1	723	3	US-08-998-416-635	Sequence 635, App
552	12.8	71.1	275	4	US-09-313-294A-3687	Sequence 3687, Ap	c 625	12.8	71.1	724	4	US-09-533-008-40	Sequence 40, Appl
553	12.8	71.1	276	4	US-09-313-294A-5760	Sequence 5760, Ap	c 626	12.8	71.1	729	4	US-09-634-238-133	Sequence 133, App
554	12.8	71.1	282	4	US-09-248-796A-1188	Sequence 1188, Ap	c 627	12.8	71.1	729	4	US-09-270-767-1796	Sequence 1796, Ap
555	12.8	71.1	284	5	US-08-238-821B-45	Sequence 45, Appl	c 628	12.8	71.1	759	4	US-09-523-029-85	Sequence 95, Appl
556	12.8	71.1	284	5	PCT-US95-05744-45	Sequence 45, Appl	c 629	12.8	71.1	759	4	US-09-270-767-3669	Sequence 3669, Ap
557	12.8	71.1	286	4	US-09-313-294A-6838	Sequence 6838, Ap	c 630	12.8	71.1	776	4	US-09-004-838-64	Sequence 64, Appl
558	12.8	71.1	288	4	US-09-313-294A-6654	Sequence 6654, Ap	c 631	12.8	71.1	784	3	US-09-270-767-18951	Sequence 2549, Ap
559	12.8	71.1	309	4	US-08-916-442-19	Sequence 19, Appl	c 632	12.8	71.1	789	4	US-09-252-991A-7444	Sequence 7444, A
560	12.8	71.1	309	4	US-09-731-924A-19	Sequence 19, Appl	c 633	12.8	71.1	792	4	US-09-248-796A-10866	Sequence 10866, A
561	12.8	71.1	309	4	US-09-317-641-19	Sequence 19, Appl	c 634	12.8	71.1	825	4	US-09-252-991A-1504	Sequence 1504, Ap
562	12.8	71.1	373	4	US-09-404-879A-190	Sequence 190, App	c 635	12.8	71.1	826	2	US-08-238-821B-51	Sequence 51, Appl
563	12.8	71.1	373	4	US-09-338-933-190	Sequence 190, App	c 636	12.8	71.1	826	2	US-08-238-821B-51	Sequence 51, Appl
564	12.8	71.1	373	4	US-09-215-681-190	Sequence 190, App	c 637	12.8	71.1	826	2	PCT-US95-05744-51	Sequence 61, Appl
565	12.8	71.1	373	4	US-09-216-003A-190	Sequence 190, App	c 638	12.8	71.1	826	5	PCT-US95-05744-51	Sequence 567, App
566	12.8	71.1	373	4	US-09-667-857-190	Sequence 190, App	c 639	12.8	71.1	864	4	US-09-248-796A-567	Sequence 87, Appl
567	12.8	71.1	375	4	US-09-404-879A-184	Sequence 184, App	c 640	12.8	71.1	879	4	US-09-110-229-1633	Sequence 1633, Ap
568	12.8	71.1	375	4	US-09-338-933-184	Sequence 184, App	c 641	12.8	71.1	882	4	US-09-707-532A-87	Sequence 701, App
569	12.8	71.1	375	4	US-09-215-681-184	Sequence 184, App	c 642	12.8	71.1	887	4	US-09-221-017B-701	Sequence 1222, Ap
570	12.8	71.1	375	4	US-09-216-003A-184	Sequence 184, App	c 643	12.8	71.1	888	3	US-08-801-740-4	Sequence 4, Appl1
571	12.8	71.1	375	4	US-09-667-857-184	Sequence 184, App	c 644	12.8	71.1	892	3	US-08-801-740-4	Sequence 4, Appl1
572	12.8	71.1	393	4	US-09-328-352-1392	Sequence 1392, Ap	c 645	12.8	71.1	892	3	US-09-134-007C-1222	Sequence 1504, Ap
573	12.8	71.1	405	4	US-09-248-796A-12580	Sequence 12580, A	c 646	12.8	71.1	896	4	US-09-107-532A-87	Sequence 1365, Ap
574	12.8	71.1	414	4	US-09-710-279-529	Sequence 529, App	c 647	12.8	71.1	936	4	US-09-252-991A-1504	Sequence 1256, Ap
575	12.8	71.1	414	4	US-09-270-767-11910	Sequence 11910, A	c 648	12.8	71.1	951	4	US-09-107-532A-87	Sequence 9489, Ap
576	12.8	71.1	462	4	US-08-916-442-20	Sequence 20, Appl	c 649	12.8	71.1	960	4	US-09-252-991A-1256	Sequence 9489, Ap
577	12.8	71.1	462	4	US-09-731-924A-20	Sequence 20, Appl	c 650	12.8	71.1	966	4	US-09-252-991A-9495	Sequence 10799, A
578	12.8	71.1	462	4	US-09-317-641-20	Sequence 20, Appl	c 651	12.8	71.1	1007	4	US-09-270-767-23881	Sequence 35, Appl
579	12.8	71.1	462	4	US-09-270-767-11694	Sequence 11694, A	c 652	12.8	71.1	1017	4	US-09-252-991A-10799	Sequence 1447, Ap
580	12.8	71.1	463	4	US-09-270-767-14424	Sequence 14424, A	c 653	12.8	71.1	1020	4	US-09-690-454-35	Sequence 2451, Ap
581	12.8	71.1	468	4	US-09-621-325-11	Sequence 11, Appl	c 654	12.8	71.1	1032	4	US-09-328-352-1652	Sequence 349, App
582	12.8	71.1	472	4	US-09-702-705-11	Sequence 11, Appl	c 655	12.8	71.1	1032	4	US-09-134-000C-2441	Sequence 23, Appl
583	12.8	71.1	472	4	US-09-736-457-11	Sequence 11, Appl	c 656	12.8	71.1	1050	1	US-08-599-171A-23	Sequence 23, Appl
584	12.8	71.1	472	4	US-09-614-124B-11	Sequence 11, Appl	c 657	12.8	71.1	1065	2	US-08-646-590B-23	Sequence 23, Appl
585	12.8	71.1	472	4	US-09-671-325-11	Sequence 11, Appl	c 658	12.8	71.1	1065	2	US-09-069-226-23	Sequence 23, Appl
586	12.8	71.1	472	4	US-09-589-184-11	Sequence 11, Appl	c 659	12.8	71.1	1065	3	US-09-412-184-23	Sequence 33, App
587	12.8	71.1	472	4	US-09-658-824-11	Sequence 11, Appl	c 660	12.8	71.1	1065	3	US-09-412-184-23	Sequence 10719, A
588	12.8	71.1	520	4	US-08-916-442-15	Sequence 15, Appl	c 661	12.8	71.1	1065	3	US-09-248-796A-333	Sequence 1744, Ap
589	12.8	71.1	520	4	US-09-317-641-15	Sequence 15, Appl	c 662	12.8	71.1	1065	4	US-09-252-991A-10719	Sequence 3754, Ap
590	12.8	71.1	520	4	US-09-916-442-12	Sequence 12, Appl	c 663	12.8	71.1	1065	4	US-09-328-352-374	Sequence 5703, Ap
591	12.8	71.1	530	4	US-08-916-442-12	Sequence 12, Appl	c 664	12.8	71.1	1066	4	US-09-448-796A-5703	Sequence 4253, Ap
592	12.8	71.1	530	4	US-09-731-924A-12	Sequence 12, Appl	c 665	12.8	71.1	1101	4	US-09-602-787A-465	Sequence 9666, Ap
593	12.8	71.1	530	4	US-09-317-641-12	Sequence 12, Appl	c 666	12.8	71.1	1107	4	US-09-328-352-374	Sequence 10719, A
594	12.8	71.1	538	4	US-09-270-767-705	Sequence 705, App	c 667	12.8	71.1	1116	4	US-09-134-000C-3188	Sequence 374, App
595	12.8	71.1	538	4	US-09-270-767-15987	Sequence 15987, A	c 668	12.8	71.1	1116	4	US-09-328-352-374	Sequence 374, App
596	12.8	71.1	544	4	US-09-058-740-1	Sequence 1, Appl1	c 669	12.8	71.1	1146	4	US-09-252-991A-10448	Sequence 10448, A
597	12.8	71.1	567	4	US-10-101-464A-189	Sequence 189, App	c 670	12.8	71.1	1173	4	US-09-328-352-374	Sequence 3754, Ap
598	12.8	71.1	576	4	US-09-621-976-16902	Sequence 16902, A	c 671	12.8	71.1	1173	4	US-09-448-796A-5703	Sequence 4253, Ap
599	12.8	71.1	586	4	US-09-270-767-14218	Sequence 14218, A	c 672	12.8	71.1	1200	4	US-09-602-787A-465	Sequence 9666, Ap
600	12.8	71.1	603	3	US-08-779-764A-2	Sequence 2, Appl1	c 673	12.8	71.1	1239	4	US-09-731-924A-1	Sequence 1, Appl1
601	12.8	71.1	603	3	US-08-779-764A-3	Sequence 3, Appl1	c 674	12.8	71.1	1243	4	US-09-731-924A-1	Sequence 1, Appl1
602	12.8	71.1	603	3	US-08-779-764A-4	Sequence 4, Appl1	c 675	12.8	71.1	1244	4	US-08-916-442-1	Sequence 1, Appl1
603	12.8	71.1	603	3	US-08-779-764A-5	Sequence 5, Appl1	c 676	12.8	71.1	1244	4	US-09-317-641-1	Sequence 37, Appl
604	12.8	71.1	603	4	US-09-563-456-2	Sequence 2, Appl1	c 677	12.8	71.1	1250	4	US-09-317-641-1	Sequence 3293, Ap
605	12.8	71.1	603	4	US-09-563-456-3	Sequence 3, Appl1	c 678	12.8	71.1	1260	4	US-09-248-796A-3393	Sequence 11058, A
606	12.8	71.1	603	4	US-09-563-456-4	Sequence 4, Appl1	c 679	12.8	71.1	1271	4	US-09-270-767-11058	Sequence 1740, Ap
607	12.8	71.1	603	4	US-09-563-456-5	Sequence 5, Appl1	c 680	12.8	71.1	1272	4	US-09-248-796A-1740	Sequence 3919, Ap
608	12.8	71.1	609	3	US-08-369-822C-3	Sequence 3, Appl1	c 681	12.8	71.1	1272	4	US-09-448-796A-3919	Sequence 3919, Ap
609	12.8	71.1	609	3	US-08-582-776C-3	Sequence 3, Appl1	c 682	12.8	71.1	1277	4	US-09-187-999-30	Sequence 30, Appl
610	12.8	71.1	609	3	US-08-434-831B-3	Sequence 3, Appl1	c 683	12.8	71.1	1284	4	US-09-540-236-26	Sequence 26, Appl
611	12.8	71.1	612	4	US-09-252-991A-1623	Sequence 1623, Ap	c 684	12.8	71.1	1284	4	US-09-540-236-26	Sequence 26, Appl

C 685	12.8	71.1	1296	4	US-09-252-991A-1670	Sequence 1670, Ap	C 758	12.8	71.1	2694	1	US-08-465-995A-3	Sequence 3, Appl1
C 686	12.8	71.1	1323	4	US-09-489-039A-5432	Sequence 5432, Ap	C 759	12.8	71.1	2694	1	US-08-465-995A-3	Sequence 3, Appl1
C 687	12.8	71.1	1371	4	US-09-350-756-3	Sequence 3, Appl1	C 760	12.8	71.1	2694	2	US-08-465-995A-3	Sequence 3, Appl1
C 688	12.8	71.1	1374	4	US-08-123-761A-2	Sequence 2, Appl1	C 761	12.8	71.1	2694	2	US-08-465-995A-3	Sequence 3, Appl1
C 689	12.8	71.1	1405	4	US-09-270-767-10838	Sequence 10838, A	C 762	12.8	71.1	2694	2	US-08-465-995A-3	Sequence 3, Appl1
C 690	12.8	71.1	1449	4	US-09-252-991A-6087	Sequence 6087, Ap	C 763	12.8	71.1	2694	2	US-08-465-995A-3	Sequence 3, Appl1
C 691	12.8	71.1	1473	4	US-08-277-031B-15	Sequence 15, Appl1	C 764	12.8	71.1	2694	1	US-08-101-593-3	Sequence 3, Appl1
C 692	12.8	71.1	1497	4	US-09-248-796A-11414	Sequence 11414, A	C 765	12.8	71.1	2763	4	US-09-774-528-12	Sequence 12, Appl1
C 693	12.8	71.1	1500	4	US-09-023-778-1	Sequence 1, Appl1	C 766	12.8	71.1	2763	4	US-09-774-528-12	Sequence 12, Appl1
C 694	12.8	71.1	1509	4	US-09-248-796A-2721	Sequence 2721, Ap	C 767	12.8	71.1	2779	4	US-08-781-966A-288	Sequence 288, Ap
C 695	12.8	71.1	1530	4	US-09-328-352-739	Sequence 739, Ap	C 768	12.8	71.1	2841	4	US-09-377-850-2	Sequence 2, Appl1
C 696	12.8	71.1	1537	4	US-09-270-767-13829	Sequence 13829, A	C 769	12.8	71.1	2859	4	US-09-248-796A-6775	Sequence 6775, Ap
C 697	12.8	71.1	1594	4	US-07-621-670-2	Sequence 14907, A	C 770	12.8	71.1	2896	4	US-09-710-279-4383	Sequence 4383, Ap
C 698	12.8	71.1	1609	4	US-09-023-655-178	Sequence 2, Appl1	C 771	12.8	71.1	2931	3	US-09-168-629-14	Sequence 14, Appl1
C 699	12.8	71.1	1657	4	US-08-941-532-5	Sequence 5, Appl1	C 772	12.8	71.1	2931	3	US-09-168-629-14	Sequence 14, Appl1
C 700	12.8	71.1	1659	4	US-08-941-532-5	Sequence 5, Appl1	C 773	12.8	71.1	3023	4	US-09-917-254-19	Sequence 19, Appl1
C 701	12.8	71.1	1659	4	US-09-270-767-12870	Sequence 12870, A	C 774	12.8	71.1	3050	4	US-09-377-850-1	Sequence 1, Appl1
C 702	12.8	71.1	1698	4	US-09-248-796A-5105	Sequence 5105, Ap	C 775	12.8	71.1	3050	4	US-09-377-850-1	Sequence 1, Appl1
C 703	12.8	71.1	1716	3	US-09-040-681A-3	Sequence 3, Appl1	C 776	12.8	71.1	3060	2	US-08-286-305A-6	Sequence 6, Appl1
C 704	12.8	71.1	1716	3	US-09-497-897-3	Sequence 3, Appl1	C 777	12.8	71.1	3060	2	US-08-441-104A-6	Sequence 6, Appl1
C 705	12.8	71.1	1746	2	US-08-201-118-2	Sequence 2, Appl1	C 778	12.8	71.1	3060	2	US-08-441-104A-6	Sequence 6, Appl1
C 706	12.8	71.1	1746	2	US-08-238-821B-2	Sequence 2, Appl1	C 779	12.8	71.1	3060	2	US-08-441-104A-6	Sequence 6, Appl1
C 707	12.8	71.1	1746	2	US-09-107-532A-3181	Sequence 3181, Ap	C 780	12.8	71.1	3111	3	US-09-417-381A-6	Sequence 6, Appl1
C 708	12.8	71.1	1746	2	US-09-023-655-1059	Sequence 1059, Ap	C 781	12.8	71.1	3132	3	US-09-221-017B-922	Sequence 922, Ap
C 709	12.8	71.1	1746	5	PCT-US95-05744-2	Sequence 262, Ap	C 782	12.8	71.1	3132	3	US-09-221-017B-922	Sequence 922, Ap
C 710	12.8	71.1	1765	4	US-09-5591-055-11	Sequence 11, Appl1	C 783	12.8	71.1	3132	3	US-09-221-017B-922	Sequence 922, Ap
C 711	12.8	71.1	1766	4	US-09-270-767-10946	Sequence 10946, A	C 784	12.8	71.1	3194	2	US-08-359-705B-1	Sequence 4, Appl1
C 712	12.8	71.1	1816	4	US-09-976-594-838	Sequence 838, Ap	C 785	12.8	71.1	3194	2	US-08-359-705B-1	Sequence 4, Appl1
C 713	12.8	71.1	1827	4	US-09-270-767-14852	Sequence 14852, A	C 786	12.8	71.1	3194	2	US-08-447-880A-1	Sequence 1, Appl1
C 714	12.8	71.1	1834	4	US-10-101-464A-262	Sequence 262, Ap	C 787	12.8	71.1	3194	3	US-08-447-880A-1	Sequence 1, Appl1
C 715	12.8	71.1	1870	2	US-08-359-705B-3	Sequence 3, Appl1	C 788	12.8	71.1	3194	3	US-08-447-880A-1	Sequence 1, Appl1
C 716	12.8	71.1	1870	2	US-08-286-846A-3	Sequence 3, Appl1	C 789	12.8	71.1	3194	3	US-08-447-880A-1	Sequence 1, Appl1
C 717	12.8	71.1	1870	2	US-08-457-880A-3	Sequence 3, Appl1	C 790	12.8	71.1	3239	3	US-10-101-464A-9	Sequence 9, Appl1
C 718	12.8	71.1	1870	2	US-08-457-880A-3	Sequence 3, Appl1	C 791	12.8	71.1	3239	3	US-10-101-464A-9	Sequence 9, Appl1
C 719	12.8	71.1	1870	3	US-08-942-562-3	Sequence 3, Appl1	C 792	12.8	71.1	3242	4	US-09-493-914-2	Sequence 2, Appl1
C 720	12.8	71.1	1870	3	US-09-156-923-3	Sequence 3, Appl1	C 793	12.8	71.1	3305	4	US-09-710-279-4236	Sequence 4236, Ap
C 721	12.8	71.1	1920	3	US-09-509-814A-3	Sequence 3, Appl1	C 794	12.8	71.1	3369	3	US-09-248-796A-2271	Sequence 2271, Ap
C 722	12.8	71.1	1920	3	US-09-134-001C-2824	Sequence 2824, Ap	C 795	12.8	71.1	3497	3	US-09-276-531-60	Sequence 60, Appl1
C 723	12.8	71.1	1920	4	US-09-920-954-3	Sequence 3, Appl1	C 796	12.8	71.1	3522	3	US-09-248-796A-5210	Sequence 5210, Ap
C 724	12.8	71.1	1935	4	US-09-270-767-11338	Sequence 11338, A	C 797	12.8	71.1	3546	3	US-09-118-442-15	Sequence 15, Appl1
C 725	12.8	71.1	1950	4	US-09-727-628-1	Sequence 6538, Ap	C 798	12.8	71.1	3546	3	US-09-118-442-15	Sequence 15, Appl1
C 726	12.8	71.1	1950	4	US-09-727-628-1	Sequence 6538, Ap	C 799	12.8	71.1	3546	3	US-09-118-442-15	Sequence 15, Appl1
C 727	12.8	71.1	1973	4	US-09-270-767-5756	Sequence 5756, Ap	C 800	12.8	71.1	3546	3	US-09-677-064-15	Sequence 15, Appl1
C 728	12.8	71.1	1973	4	US-09-270-767-21038	Sequence 21038, A	C 801	12.8	71.1	3617	4	US-10-140-002-09	Sequence 409, Ap
C 729	12.8	71.1	1981	4	US-09-620-312D-715	Sequence 715, Ap	C 802	12.8	71.1	3693	4	US-09-976-594-808	Sequence 908, Ap
C 730	12.8	71.1	2019	4	US-09-270-767-14598	Sequence 14598, A	C 803	12.8	71.1	3693	4	US-09-976-594-808	Sequence 908, Ap
C 731	12.8	71.1	2061	4	US-09-252-991A-1544	Sequence 1544, A	C 804	12.8	71.1	3707	1	US-08-271-454-1	Sequence 359, Ap
C 732	12.8	71.1	2092	4	US-09-468-253B-13	Sequence 13, Appl1	C 805	12.8	71.1	3707	5	PCT-US95-08180-1	Sequence 1, Appl1
C 733	12.8	71.1	2100	4	US-09-270-767-11711	Sequence 11711, A	C 806	12.8	71.1	3966	3	US-09-215-131-1	Sequence 1, Appl1
C 734	12.8	71.1	2106	1	US-07-718-535-4	Sequence 4, Appl1	C 807	12.8	71.1	3966	3	US-09-215-131-1	Sequence 1, Appl1
C 735	12.8	71.1	2106	1	US-08-161-999-4	Sequence 4, Appl1	C 808	12.8	71.1	4020	3	US-09-050-159-130	Sequence 130, Ap
C 736	12.8	71.1	2122	4	US-09-367-895-38	Sequence 38, Appl1	C 809	12.8	71.1	4020	3	US-09-050-159-130	Sequence 130, Ap
C 737	12.8	71.1	2190	4	US-09-023-655-117	Sequence 117, Ap	C 810	12.8	71.1	4024	3	US-09-162-884-18	Sequence 18, Appl1
C 738	12.8	71.1	2268	2	US-08-880-853-1	Sequence 1, Appl1	C 811	12.8	71.1	4179	4	US-09-710-279-3949	Sequence 3949, Ap
C 739	12.8	71.1	2268	2	US-09-039-125A-1	Sequence 1, Appl1	C 812	12.8	71.1	4854	4	US-08-961-527-178	Sequence 178, Ap
C 740	12.8	71.1	2268	2	US-09-099-124A-1	Sequence 1, Appl1	C 813	12.8	71.1	4970	4	US-09-816-095-1	Sequence 1, Appl1
C 741	12.8	71.1	2268	2	US-09-197-008-1	Sequence 1, Appl1	C 814	12.8	71.1	5058	3	US-08-934-386-8	Sequence 8, Appl1
C 742	12.8	71.1	2268	2	US-09-032-476-1	Sequence 1, Appl1	C 815	12.8	71.1	5198	3	US-08-133-761A-1	Sequence 1, Appl1
C 743	12.8	71.1	2268	3	US-08-890-854-1	Sequence 1, Appl1	C 816	12.8	71.1	5386	4	US-09-535-008-60	Sequence 60, Appl1
C 744	12.8	71.1	2268	3	US-09-023-324-1	Sequence 1, Appl1	C 817	12.8	71.1	5468	4	US-09-535-008-60	Sequence 60, Appl1
C 745	12.8	71.1	2268	4	US-09-109-886-1	Sequence 1, Appl1	C 818	12.8	71.1	5471	4	US-09-535-008-60	Sequence 60, Appl1
C 746	12.8	71.1	2271	3	US-08-910-820-8	Sequence 8, Appl1	C 819	12.8	71.1	5471	4	US-09-535-008-60	Sequence 60, Appl1
C 747	12.8	71.1	2271	3	US-09-844-908-8	Sequence 8, Appl1	C 820	12.8	71.1	5471	4	US-09-535-008-60	Sequence 60, Appl1
C 748	12.8	71.1	2309	4	US-09-367-895-39	Sequence 39, Appl1	C 821	12.8	71.1	5480	4	US-09-535-008-60	Sequence 60, Appl1
C 749	12.8	71.1	2463	4	US-08-339-578-1	Sequence 39, Appl1	C 822	12.8	71.1	5544	4	US-09-561-818A-11	Sequence 11, Appl1
C 750	12.8	71.1	2478	1	US-08-481-625-1	Sequence 1, Appl1	C 823	12.8	71.1	5564	4	US-09-535-008-60	Sequence 60, Appl1
C 751	12.8	71.1	2592	4	US-09-321-017B-642	Sequence 642, Ap	C 824	12.8	71.1	5567	4	US-09-535-008-60	Sequence 60, Appl1
C 752	12.8	71.1	2609	4	US-09-437-568A-1	Sequence 1, Appl1	C 825	12.8	71.1	5573	4	US-09-535-008-60	Sequence 60, Appl1
C 753	12.8	71.1	2631	4	US-09-270-767-10462	Sequence 10462, A	C 826	12.8	71.1	5576	4	US-09-535-008-60	Sequence 60, Appl1
C 754	12.8	71.1	2658	3	US-08-369-822C-20	Sequence 20, Appl1	C 827	12.8	71.1	5566	3	US-09-535-008-60	Sequence 60, Appl1
C 755	12.8	71.1	2658	3	US-08-582-776C-33	Sequence 33, Appl1	C 828	12.8	71.1	5824	4	US-08-938-105-2	Sequence 2, Appl1
C 756	12.8	71.1	2658	3	US-08-434-831B-58	Sequence 58, Appl1	C 829	12.8	71.1	6326	4	US-09-561-818A-58	Sequence 58, Appl1
C 757	12.8	71.1	2694	1	US-08-465-995A-1	Sequence 1, Appl1	C 830	12.8	71.1	6701	3	US-09-004-838-107	Sequence 107, Ap

831	12.8	71.1	7454	4	US-08-836-687B-29	Sequence 29, Appl	904	12.4	68.9	515	4	US-09-621-976-1563	Sequence 1563, Ap
832	12.8	71.1	8424	4	US-09-543-681A-108	Sequence 808, Appl	905	12.4	68.9	519	4	US-09-621-976-10460	Sequence 10460, A
833	12.8	71.1	8910	3	US-08-369-822C-19	Sequence 18, Appl	906	12.4	68.9	522	4	US-09-621-976-1564	Sequence 1564, A
834	12.8	71.1	8910	3	US-08-779-764A-1	Sequence 1, Appl	907	12.4	68.9	537	4	US-09-248-796A-1162	Sequence 1162, Ap
835	12.8	71.1	8910	3	US-08-589-776C-19	Sequence 19, Appl	908	12.4	68.9	540	3	US-08-714-918-56	Sequence 56, Appl
836	12.8	71.1	8910	3	US-08-434-831B-19	Sequence 19, Appl	909	12.4	68.9	540	3	US-09-265-315-56	Sequence 56, Appl
837	12.8	71.1	8910	3	US-09-563-456-1	Sequence 1, Appl	910	12.4	68.9	540	3	US-09-265-315-56	Sequence 56, Appl
838	12.8	71.1	8910	3	US-09-563-456-1	Sequence 1, Appl	911	12.4	68.9	540	3	US-09-265-315-56	Sequence 56, Appl
839	12.8	71.1	10470	4	US-08-956-171E-20	Sequence 20, Appl	912	12.4	68.9	540	3	US-09-265-315-56	Sequence 56, Appl
840	12.8	71.1	10470	4	US-08-781-986A-20	Sequence 20, Appl	913	12.4	68.9	540	3	US-09-265-315-56	Sequence 56, Appl
841	12.8	71.1	11014	4	US-08-956-171E-91	Sequence 91, Appl	914	12.4	68.9	540	3	US-09-265-315-56	Sequence 56, Appl
842	12.8	71.1	11580	3	US-08-781-986A-91	Sequence 91, Appl	915	12.4	68.9	546	3	US-09-134-001C-55	Sequence 55, Appl
843	12.8	71.1	14113	3	US-09-223-134-1	Sequence 4, Appl	916	12.4	68.9	549	4	US-09-352-991A-3125	Sequence 3125, Ap
844	12.8	71.1	14113	3	US-08-992-801-1	Sequence 1, Appl	917	12.4	68.9	549	4	US-09-352-991A-3125	Sequence 3125, Ap
845	12.8	71.1	14113	3	US-09-223-535-1	Sequence 1, Appl	918	12.4	68.9	549	4	US-09-352-991A-3125	Sequence 3125, Ap
846	12.8	71.1	17656	4	US-09-433-579-3	Sequence 3, Appl	919	12.4	68.9	549	4	US-09-352-991A-3125	Sequence 3125, Ap
847	12.8	71.1	141171	4	US-08-311-731A-122	Sequence 122, App	920	12.4	68.9	602	1	US-08-229-393-1	Sequence 1, Appl
848	12.8	71.1	74962	4	US-09-685-853A-3	Sequence 3, Appl	921	12.4	68.9	606	4	US-09-489-039A-4415	Sequence 4415, Ap
849	12.8	71.1	99916	4	US-09-815-055-3	Sequence 3, Appl	922	12.4	68.9	618	4	US-09-669-751-139	Sequence 139, Appl
850	12.8	71.1	129908	4	US-09-585-858-1	Sequence 1, Appl	923	12.4	68.9	618	4	US-09-248-796A-3806	Sequence 3806, Ap
851	12.8	71.1	169998	4	US-09-676-610B-34	Sequence 24, Appl	924	12.4	68.9	637	4	US-09-110-279-693	Sequence 693, Appl
852	12.8	71.1	197496	4	US-09-877-177A-10	Sequence 10, Appl	925	12.4	68.9	637	4	US-09-110-279-693	Sequence 693, Ap
853	12.8	71.1	269223	4	US-09-596-002-41	Sequence 41, Appl	926	12.4	68.9	633	3	US-09-134-001C-578	Sequence 578, Appl
854	12.8	71.1	580073	4	US-08-545-558D-41	Sequence 41, Appl	927	12.4	68.9	633	4	US-09-248-796A-906	Sequence 906, Appl
855	12.8	71.1	640681	4	US-09-790-988-1	Sequence 1, Appl	928	12.4	68.9	643	4	US-09-621-976-578	Sequence 578, Appl
856	12.8	71.1	1230025	4	US-09-198-452A-1	Sequence 1, Appl	929	12.4	68.9	657	4	US-09-428-796A-1648	Sequence 1648, Ap
857	12.4	68.9	20	4	US-09-198-452A-5994	Sequence 5994, Ap	930	12.4	68.9	658	4	US-09-799-451-697	Sequence 697, Appl
858	12.4	68.9	65	3	US-09-191-852-17	Sequence 17, Appl	931	12.4	68.9	660	4	US-09-248-796A-4480	Sequence 4480, Ap
859	12.4	68.9	65	3	US-08-817-906-17	Sequence 17, Appl	932	12.4	68.9	666	4	US-09-328-352-589	Sequence 589, Appl
860	12.4	68.9	65	5	PCT-UB95-13376-17	Sequence 17, Appl	933	12.4	68.9	667	4	US-09-270-767-14063	Sequence 14063, A
861	12.4	68.9	87	4	US-09-513-999C-13601	Sequence 13601, A	934	12.4	68.9	678	3	US-09-232-479-19	Sequence 19, Appl
862	12.4	68.9	189	4	US-09-248-796A-11625	Sequence 11625, A	935	12.4	68.9	678	4	US-09-248-796A-1620	Sequence 1620, Ap
863	12.4	68.9	220	4	US-09-513-999C-10749	Sequence 10749, A	936	12.4	68.9	678	4	US-09-248-796A-1620	Sequence 1620, Ap
864	12.4	68.9	221	1	US-08-606-789-12	Sequence 12, Appl	937	12.4	68.9	667	3	US-09-453-702B-35	Sequence 35, Appl
865	12.4	68.9	221	2	US-09-111-348-12	Sequence 12, Appl	938	12.4	68.9	666	4	US-09-252-991A-1137	Sequence 1137, A
866	12.4	68.9	255	4	US-09-248-796A-1153	Sequence 1153, Ap	939	12.4	68.9	701	4	US-09-641-638-179	Sequence 179, Appl
867	12.4	68.9	279	4	US-09-248-796A-8913	Sequence 8913, Ap	940	12.4	68.9	701	4	US-10-170-097-179	Sequence 179, Appl
868	12.4	68.9	286	4	US-09-313-294A-4525	Sequence 4525, Ap	941	12.4	68.9	705	4	US-09-328-352-860	Sequence 860, Appl
869	12.4	68.9	286	4	US-09-513-999C-14678	Sequence 14678, A	942	12.4	68.9	708	4	US-09-252-991A-3024	Sequence 3024, Ap
870	12.4	68.9	291	4	US-09-248-796A-5206	Sequence 5206, Ap	943	12.4	68.9	711	4	US-09-252-991A-3024	Sequence 3024, Ap
871	12.4	68.9	292	4	US-09-313-294A-5755	Sequence 5755, Ap	944	12.4	68.9	711	4	US-09-248-796A-4748	Sequence 4748, Ap
872	12.4	68.9	306	4	US-09-313-294A-5012	Sequence 5012, A	945	12.4	68.9	717	4	US-09-107-532A-1765	Sequence 1765, Ap
873	12.4	68.9	311	4	US-09-513-999C-15612	Sequence 15612, A	946	12.4	68.9	792	4	US-09-248-796A-878	Sequence 878, Appl
874	12.4	68.9	312	3	US-09-191-852-20	Sequence 20, Appl	947	12.4	68.9	795	4	US-09-248-796A-878	Sequence 878, Appl
875	12.4	68.9	312	3	US-08-817-906-20	Sequence 20, Appl	948	12.4	68.9	818	4	US-08-956-171E-742	Sequence 742, Appl
876	12.4	68.9	312	5	PCT-UB95-13376-20	Sequence 20, Appl	949	12.4	68.9	818	4	US-08-781-986A-742	Sequence 742, Appl
877	12.4	68.9	332	4	US-09-270-767-31505	Sequence 31505, A	950	12.4	68.9	833	4	US-09-252-991A-1378	Sequence 1378, A
878	12.4	68.9	336	4	US-09-489-039A-6517	Sequence 6517, Ap	951	12.4	68.9	837	4	US-09-583-110-282	Sequence 282, Appl
879	12.4	68.9	351	3	US-09-134-001C-2503	Sequence 2503, Ap	952	12.4	68.9	882	4	US-09-710-279-2645	Sequence 2645, Ap
880	12.4	68.9	366	4	US-09-248-796A-2185	Sequence 2185, Ap	953	12.4	68.9	890	4	US-09-270-767-1518	Sequence 1518, A
881	12.4	68.9	367	4	US-09-702-705-945	Sequence 945, App	954	12.4	68.9	890	4	US-09-328-352-1956	Sequence 1956, A
882	12.4	68.9	367	4	US-09-736-457-945	Sequence 945, App	955	12.4	68.9	897	5	US-09-248-796A-3605	Sequence 3605, Ap
883	12.4	68.9	367	4	US-09-614-124B-945	Sequence 945, App	956	12.4	68.9	900	4	US-09-248-796A-6667	Sequence 6667, Ap
884	12.4	68.9	367	4	US-09-671-325-945	Sequence 945, App	957	12.4	68.9	915	4	US-09-248-796A-710	Sequence 710, Appl
885	12.4	68.9	367	4	US-09-658-824-945	Sequence 945, App	958	12.4	68.9	918	4	US-09-248-796A-6594	Sequence 6594, Ap
886	12.4	68.9	375	4	US-09-513-999C-322	Sequence 322, App	959	12.4	68.9	921	4	US-09-270-767-1529	Sequence 1529, A
887	12.4	68.9	380	4	US-09-513-999C-9573	Sequence 9573, Ap	960	12.4	68.9	921	4	US-09-270-767-1611	Sequence 1611, A
888	12.4	68.9	385	4	US-09-270-767-1425	Sequence 1425, Ap	961	12.4	68.9	945	4	US-09-328-352-1956	Sequence 1956, A
889	12.4	68.9	385	4	US-09-270-767-16707	Sequence 16707, A	962	12.4	68.9	945	4	US-09-248-796A-3605	Sequence 3605, Ap
890	12.4	68.9	392	4	US-09-621-976-15659	Sequence 15659, A	963	12.4	68.9	975	4	US-09-252-991A-5535	Sequence 5535, Ap
891	12.4	68.9	420	4	US-09-248-796A-5722	Sequence 5722, Ap	964	12.4	68.9	1001	4	US-09-641-638-180	Sequence 180, Appl
892	12.4	68.9	435	3	US-09-134-001C-123	Sequence 123, App	965	12.4	68.9	1001	4	US-10-170-097-180	Sequence 180, Appl
893	12.4	68.9	440	4	US-09-513-999C-34551	Sequence 34591, A	966	12.4	68.9	1005	4	US-09-248-796A-2263	Sequence 2263, Ap
894	12.4	68.9	441	4	US-09-252-991A-13884	Sequence 13884, A	967	12.4	68.9	1026	4	US-09-107-532A-3333	Sequence 3333, Ap
895	12.4	68.9	459	4	US-09-270-767-7119	Sequence 7119, Ap	968	12.4	68.9	1032	4	US-08-956-171E-247	Sequence 247, Appl
896	12.4	68.9	459	4	US-09-270-767-22401	Sequence 22401, A	969	12.4	68.9	1032	4	US-08-781-986A-247	Sequence 247, Appl
897	12.4	68.9	462	4	US-09-513-999C-22898	Sequence 22848, A	970	12.4	68.9	1068	4	US-09-107-532A-163	Sequence 163, Appl
898	12.4	68.9	464	4	US-09-248-796A-3653	Sequence 3653, Ap	971	12.4	68.9	1068	4	US-09-270-767-2557	Sequence 2557, Ap
899	12.4	68.9	474	4	US-09-621-976-1566	Sequence 1566, Ap	972	12.4	68.9	1074	4	US-09-270-767-15180	Sequence 15180, A
900	12.4	68.9	477	4	US-09-248-796A-5721	Sequence 5721, Ap	973	12.4	68.9	1101	4	US-09-248-796A-863	Sequence 863, Appl
901	12.4	68.9	478	4	US-09-621-976-10037	Sequence 10037, A	974	12.4	68.9	1107	4	US-08-956-171E-638	Sequence 638, Appl
902	12.4	68.9	483	4	US-09-583-110-2569	Sequence 2569, Ap	975	12.4	68.9	1107	4	US-08-781-986A-658	Sequence 658, Appl
903	12.4	68.9	507	4	US-09-248-796A-7370	Sequence 7370, Ap	976	12.4	68.9	1110	4	US-09-248-796A-1340	Sequence 1340, Ap


```
977 12.4 68.9 1155 4 US-09-252-991A-13685 Sequence 13685, A
978 12.4 68.9 1158 4 US-09-248-796A-2478 Sequence 2478, Ap
979 12.4 68.9 1164 3 US-09-134-001C-1461 Sequence 1461, Ap
980 12.4 68.9 1170 4 US-09-248-796A-3621 Sequence 3621, Ap
981 12.4 68.9 1182 4 US-09-248-796A-4758 Sequence 4758, Ap
982 12.4 68.9 1183 4 US-09-684-385-9 Sequence 9, Appl1
983 12.4 68.9 1203 4 US-09-248-796A-3748 Sequence 3748, Ap
984 12.4 68.9 1205 2 US-08-680-385-7 Sequence 7, Appl1
985 12.4 68.9 1214 4 US-09-506-286B-15 Sequence 15, Appl
986 12.4 68.9 1214 4 US-09-506-286B-18 Sequence 18, Appl
987 12.4 68.9 1214 4 US-09-762-861B-15 Sequence 15, Appl
988 12.4 68.9 1214 4 US-09-762-861B-18 Sequence 18, Appl
989 12.4 68.9 1214 4 US-10-065-133A-15 Sequence 15, Appl
990 12.4 68.9 1214 4 US-10-065-133A-18 Sequence 18, Appl
991 12.4 68.9 1215 4 US-09-328-352-3884 Sequence 3884, Ap
992 12.4 68.9 1229 1 US-08-180-209B-54 Sequence 54, Appl
993 12.4 68.9 1229 3 US-08-474-853-54 Sequence 54, Appl
994 12.4 68.9 1229 3 US-09-166-205B-54 Sequence 54, Appl
995 12.4 68.9 1229 5 PCT-US94-02629-54 Sequence 54, Appl
996 12.4 68.9 1241 4 US-09-506-286B-13 Sequence 13, Appl
997 12.4 68.9 1241 4 US-09-506-286B-16 Sequence 16, Appl
998 12.4 68.9 1241 4 US-09-762-861B-13 Sequence 13, Appl
999 12.4 68.9 1241 4 US-09-762-861B-16 Sequence 16, Appl
1000 12.4 68.9 1241 4 US-10-065-133A-13 Sequence 13, Appl
```

ALIGNMENTS

```
RESULT 1
US-09-134-000C-2523
; Sequence 2523, Application US/09134000C
; Patent No. 6617156
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; FILE OF INVENTION: ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 032796-032
; CURRENT APPLICATION NUMBER: US/09/134, 000C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/055, 778
; PRIOR FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2523
; LENGTH: 534
; TYPE: DNA
; ORGANISM: Enterococcus faecalis
US-09-134-000C-2523
```

```
Query Match 85.6%; Score 15.4; DB 4; Length 534;
Best Local Similarity 94.1%; Pred. No. 94;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
OY 2 AGACACACAGATTGCG 18
Db 465 AGACACACAGATTGCG 481
```

```
RESULT 2
US-08-968-563-7
; Sequence 7, Application US/08968563
```

```
; Patent No. 6013494
; GENERAL INFORMATION:
; APPLICANT: CHARLES E. NAKAMURA
; APPLICANT: ANTHONY A. GATENBY
; APPLICANT: AMY (KUANG-HUA) HSU
; APPLICANT: RICHARD D. LA REAU
; APPLICANT: SHARON L. HAYNIE
; APPLICANT: MARIA DIAZ-TORRES
; APPLICANT: DONALD E. TRIMBUR
; APPLICANT: GREGORY W. WHITED
; APPLICANT: VASANTHA NAGARAJAN
```

```
APPLICANT: MARK S. PAYNE
APPLICANT: STEPHEN K. PICATAGGIO
APPLICANT: RAMESCH V. NAIR
TITLE OF INVENTION: METHOD FOR THE RECOMBINANT
TITLE OF INVENTION: PRODUCTION OF 1,3-PROPANEDIOL
NUMBER OF SEQUENCES: 49
CORRESPONDENCE ADDRESS:
ADDRESSER: E. I. DU PONT DE NEMOURS AND COMPANY
STREET: 1007 MARKET STREET
CITY: WILMINGTON
STATE: DELAWARE
COUNTRY: U.S.A.
ZIP: 19898
ADDRESSER: GENENCOR INTERNATIONAL, INC.
STREET: 4 CAMBRIDGE PLACE
STREET: 1870 SOUTH WINTON ROAD
CITY: ROCHESTER
STATE: NEW YORK
COUNTRY: U.S.A.
```

```
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.50 INCH DISKETTE
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: MICROSOFT WORD FOR WINDOWS 95
SOFTWARE: MICROSOFT WORD VERSION 7.0A
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/968,563
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/030,601
FILING DATE: NOVEMBER 13, 1996
ATTORNEY/AGENT INFORMATION:
NAME: FLOYD, LINDA AXAMETHY
REGISTRATION NUMBER: 33,692
REFERENCE/DOCKET NUMBER: CR-9982
TELECOMMUNICATION INFORMATION:
TELEPHONE: 302-892-8112
TELEFAX: 302-773-0164
```

```
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 3178 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
ORIGINAL SOURCE:
ORGANISM: GUT2
US-08-968-563-7
```

```
Query Match 85.6%; Score 15.4; DB 3; Length 3178;
Best Local Similarity 94.1%; Pred. No. 1.2e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
OY 2 AGACACACAGATTGCG 18
Db 2154 AGACACACAGATTGCG 2170
```

```
RESULT 3
US-08-969-683A-7
; Sequence 7, Application US/08969683A
```

```
; Patent No. 6136576
; GENERAL INFORMATION:
; APPLICANT: GENENCOR INTERNATIONAL, INC.
; TITLE OF INVENTION: METHOD FOR THE RECOMBINANT
; FILE OF INVENTION: PRODUCTION OF 1,3 PROPANEDIOL
; NUMBER OF SEQUENCES: 68
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Genencor International, Inc.
; STREET: 4 Cambridge Place
; STREET: 1870 South Winton Road
; CITY: Rochester
; STATE: NY
```


COUNTRY: U.S.A
ZIP: 14618
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows
SOFTWARE: FastSeq for Windows Version 2.0b
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/969,683A
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US97/20873
FILING DATE: 13-NOV-1997
APPLICATION NUMBER: 60/030,601
FILING DATE: 13-NOV-1996
ATTORNEY/AGENT INFORMATION:
NAME: Gialster, Debra
REGISTRATION NUMBER: 33,888
REFERENCE/DOCKET NUMBER: GC 369-2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-864-7620
TELEFAX: 650-845-6504
TELEX:
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 3178 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
ORIGINAL SOURCE:
ORGANISM: GUT2
US-08-969-683A-7

Query Match 85.6%; Score 15.4; DB 3; Length 3178;
Best Local Similarity 94.1%; Pred. No. 1.2e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 AGAACAACAAGATTGGC 18
DB 2154 AGAACAACAAGATTGGC 2170

RESULT 4
US-09-297-928-3
Sequence 3, Application US/09297928
Patent No. 6358716
GENERAL INFORMATION:
APPLICANT: BULTHUIS, BEN A.
GATENBY, ANTHONY A.
HAYNIE, SHARON L.
HSU, AMY K.
LARBAU, RICHARD D.
TITLE OF INVENTION: METHOD FOR THE PRODUCTION OF
GLYCEROL BY RECOMBINANT
ORGANISMS
NUMBER OF SEQUENCES: 25
CORRESPONDENCE ADDRESS:
ADDRESSEE: GENENCOR INTERNATIONAL, INC.
STREET: 4 CAMBRIDGE PLACE
CITY: ROCHESTER
STATE: NEW YORK
COUNTRY: U.S.A.
ZIP: 14618
COMPUTER READABLE FORM:
MEDIUM TYPE: DISKETTE, 3.5 INCH
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: MICROSOFT WINDOWS 95
SOFTWARE: MICROSOFT WORD VERSION 7.0A
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/297,928

FILING DATE: 11-May-1999
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/03602
FILING DATE: NOVEMBER 13, 1996
ATTORNEY/AGENT INFORMATION:
NAME: FLOYD, LINDA AXAMETHY
REGISTRATION NUMBER: 33,692
REFERENCE/DOCKET NUMBER: CR-9981-P1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 302-892-8112
TELEFAX: 302-773-0164
TELEX: 6717325
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 3178 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-09-297-928-3

Query Match 85.6%; Score 15.4; DB 3; Length 3178;
Best Local Similarity 94.1%; Pred. No. 1.2e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 AGAACAACAAGATTGGC 18
DB 2154 AGAACAACAAGATTGGC 2170

RESULT 5
US-09-276-599-4/c
Sequence 4, Application US/09276599
Patent No. 6380459
GENERAL INFORMATION:
APPLICANT: Pereira, J. Ranjan
TITLE OF INVENTION: Composition and methods for the
TITLE OF INVENTION: modification of gene expression
FILE REFERENCE: 11000,1036
CURRENT APPLICATION NUMBER: US/09/276,599
NUMBER OF SEQ ID NOS: 21
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 4
LENGTH: 485
TYPE: DNA
ORGANISM: Pinus radiata
FEATURE:
NAME/KEY: 5'UTR
LOCATION: (1)...(431)
NAME/KEY: TATA signal
LOCATION: (350)...(356)
NAME/KEY: CAAT signal
LOCATION: (326)...(333)
US-09-276-599-4

Query Match 83.3%; Score 15; DB 3; Length 485;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 AACACAAGATTGGC 18
DB 409 AACACAAGATTGGC 395

RESULT 6
US-09-598-401c-4/c
Sequence 4, Application US/09598401C
Patent No. 6596925
GENERAL INFORMATION:

```
/ APPLICANT: Ferreira, J. Ranjan
/ APPLICANT: Sagleton, Clare
/ APPLICANT: Rice, Stephen J.
/ TITLE OF INVENTION: Compositions and Methods for the
/ TITLE OF INVENTION: Modification of Gene Expression
/ FILE REFERENCE: 11000.10362
/ CURRENT APPLICATION NUMBER: US/09/598,401C
/ PRIOR FILING DATE: 2000-06-20
/ PRIOR APPLICATION NUMBER: U.S. No. 6596925 60/146,591
/ PRIOR FILING DATE: 1999-07-30
/ PRIOR APPLICATION NUMBER: PCT/NZ00/00018
/ PRIOR FILING DATE: 2000-02-24
/ PRIOR APPLICATION NUMBER: U.S. No. 6596925 09/276,599
/ NUMBER OF SEQ ID NOS: 120
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 4
/ LENGTH: 485
/ TYPE: DNA
/ ORGANISM: Pinus radiata
/ FEATURE:
/ NAME/KEY: 5'UTR
/ LOCATION: (1)...(431)
/ FEATURE:
/ NAME/KEY: TATA signal
/ LOCATION: (350)...(356)
/ FEATURE:
/ NAME/KEY: CAT signal
/ LOCATION: (326)...(333)
/ US-09-598-401C-4
```

Query Match 83.3%; Score 15; DB 4; Length 485;

Best Local Similarity 100.0%; Pred. No. 1.5e+02; Indels 0; Gaps 0;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 4 AACACACAGATTGCG 18
DB 409 AACACACAGATTGCG 395
```

```
RESULT 7
US-09-248-796A-6542
/ Sequence 6542, Application US/09248796A
/ Patent No. 6747137
/ GENERAL INFORMATION:
/ APPLICANT: Keith Weinstein et al
/ TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
/ FILE REFERENCE: 107196.132
/ CURRENT APPLICATION NUMBER: US/09/248,796A
/ PRIOR FILING DATE: 1999-02-12
/ PRIOR APPLICATION NUMBER: US 60/074,725
/ PRIOR FILING DATE: 1998-02-13
/ PRIOR APPLICATION NUMBER: US 60/096,409
/ PRIOR FILING DATE: 1998-08-13
/ NUMBER OF SEQ ID NOS: 28208
/ SEQ ID NO 6542
/ LENGTH: 1359
/ TYPE: DNA
/ ORGANISM: Candida albicans
/ US-09-248-796A-6542
```

Query Match 83.3%; Score 15; DB 4; Length 1359;

Best Local Similarity 100.0%; Pred. No. 1.7e+02;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 2 AGAACACACAGATTG 16
DB 696 AGAACACACAGATTG 710
```

RESULT 8
US-09-495-050A-15

```
/ Sequence 15, Application US/09495050A
/ Patent No. 6492505
/ GENERAL INFORMATION:
/ APPLICANT: Roopa, Reddy
/ APPLICANT: Guegler, Karl, J.
/ APPLICANT: Au-Toung, Janice
/ TITLE OF INVENTION: COMPOSITION FOR DETECTION OF GENES ENCODING MEMBRANE-ASSOCIATED PR
/ FILE REFERENCE: PA-0013 US
/ CURRENT APPLICATION NUMBER: US/09/495,050A
/ PRIOR FILING DATE: 2000-01-31
/ PRIOR APPLICATION NUMBER: 60/118,318
/ PRIOR FILING DATE: February 1, 1999
/ NUMBER OF SEQ ID NOS: 305
/ SOFTWARE: PERL Program
/ SEQ ID NO 15
/ LENGTH: 676
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: misc feature
/ OTHER INFORMATION: Incyte ID No. 6492505 227799CBI
/ US-09-495-050A-15
```

Query Match 82.2%; Score 14.8; DB 4; Length 676;

Best Local Similarity 88.9%; Pred. No. 1.9e+02;

Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
QY 1 GAGAACACACAGATTGCG 18
DB 30 GAGAACACACAGATTGCG 47
```

```
RESULT 9
US-09-270-767-10453
/ Sequence 10453, Application US/09270767
/ Patent No. 6703491
/ GENERAL INFORMATION:
/ APPLICANT: Homburger et al.
/ TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
/ FILE REFERENCE: File Reference: 7326-094
/ CURRENT APPLICATION NUMBER: US/09/270,767
/ PRIOR FILING DATE: 1999-03-17
/ NUMBER OF SEQ ID NOS: 62517
/ SOFTWARE: Patent In Ver. 2.0
/ SEQ ID NO 10453
/ LENGTH: 1098
/ TYPE: DNA
/ ORGANISM: Drosophila melanogaster
/ US-09-270-767-10453
```

Query Match 82.2%; Score 14.8; DB 4; Length 1098;

Best Local Similarity 88.9%; Pred. No. 2.1e+02;

Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
QY 1 GAGAACACACAGATTGCG 18
DB 660 GAGAACACACAGATTGCG 677
```

```
RESULT 10
US-09-328-352-3042/C
/ Sequence 3042, Application US/09328352
/ Patent No. 6562958
/ GENERAL INFORMATION:
/ APPLICANT: Gary L. Breton et al.
/ TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
/ FILE REFERENCE: GTC99-03PA
/ CURRENT APPLICATION NUMBER: US/09/328,352
/ PRIOR FILING DATE: 1999-06-04
/ NUMBER OF SEQ ID NOS: 8252
/ SEQ ID NO 3042
/ LENGTH: 2727
```

TYPE: DNA
ORGANISM: Acinetobacter baumannii
US-09-328-352-3042

Query Match 82.2%; Score 14.8; DB 4; Length 2727;
Best Local Similarity 88.9%; Pred. No. 2.3e+02;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GAGAACACACAGATTGCG 18
DB 2114 GAGAACACACATGATTGCG 2097

RESULT 11

US-08-804-227C-7
Sequence 7, Application US/08804227C
Patent No. 5876991

GENERAL INFORMATION:

APPLICANT: Dehoff, Bradley S.
APPLICANT: Kuhstoss, Stuart A.
APPLICANT: Rostock, Paul R., Jr.
APPLICANT: Sutton, Kimberly L.
TITLE OF INVENTION: POLYKETIDE SYNTHASE GENES
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:

ADDRESSEE: THOMAS G. PLANT 1501
STREET: LILLY CORPORATE CENTER
CITY: INDIANAPOLIS
STATE: IN
COUNTRY: USA
ZIP: 46285

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM Compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: ASCII(DOS) Text only
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/804.227C
FILING DATE: February 21, 1997

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Plant, Thomas, G.
REGISTRATION NUMBER: 35,784
REFERENCE/DOCKET NUMBER: X-8231
TELECOMMUNICATION INFORMATION:
TELEPHONE: 317-276-2459
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:

LENGTH: 4437 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)

FEATURE:

NAME/KEY: CDS
LOCATION: 350..14002

FEATURE:

NAME/KEY: CDS
LOCATION: 14046..20036

FEATURE:

NAME/KEY: CDS
LOCATION: 20110..31284

FEATURE:

NAME/KEY: CDS
LOCATION: 31329..36071

FEATURE:

NAME/KEY: CDS
LOCATION: 36155..41830

US-08-804-227C-7

Query Match 82.2%; Score 14.8; DB 2; Length 44377;
Best Local Similarity 88.9%; Pred. No. 3.2e+02;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GAGAACACACAGATTGCG 18
DB 14055 GAGAACACACAGATCCGC 14072

US-08-804-198-1
Sequence 1, Application US/08804198
Patent No. 5945320

GENERAL INFORMATION:

APPLICANT: BURGESS, Stanley G.
APPLICANT: Kuhstoss, Stuart A.
APPLICANT: Rao, Nagaraja R.
APPLICANT: Richards, Mark A.
APPLICANT: Rostock, Paul R., Jr.
TITLE OF INVENTION: PLATENOILIDE SYNTHASE GENE
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:

ADDRESSEE: PAUL R. CANTRELL 1138
STREET: LILLY CORPORATE CENTER
CITY: INDIANAPOLIS
STATE: IN
COUNTRY: USA
ZIP: 46285

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: Macintosh
OPERATING SYSTEM: Macintosh 7.0
SOFTWARE: Microsoft Word 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/804.198
FILING DATE:

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: CANTRELL, PAUL R.
REGISTRATION NUMBER: 36,470
REFERENCE/DOCKET NUMBER: P9113
TELECOMMUNICATION INFORMATION:
TELEPHONE: 317-276-3885
INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 44377 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)

FEATURE:

NAME/KEY: CDS
LOCATION: 350..14002

FEATURE:

NAME/KEY: CDS
LOCATION: 14046..20036

FEATURE:

NAME/KEY: CDS
LOCATION: 20110..31284

FEATURE:

NAME/KEY: CDS
LOCATION: 31329..36071

FEATURE:

NAME/KEY: CDS
LOCATION: 36155..41830

US-08-804-198-1

Query Match 82.2%; Score 14.8; DB 2; Length 44377;
Best Local Similarity 88.9%; Pred. No. 3.2e+02;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GAGAACACACAGATTGCG 18
DB 14055 GAGAACACACAGATCCGC 14072

RESULT 13
US-09-103-840A-2
; Sequence 2, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 4403765
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; FEATURE:
; OTHER INFORMATION: CDC 1551
; OTHER INFORMATION: "n" bases at various positions throughout the sequence
; OTHER INFORMATION: represent a, t, c or g
US-09-103-840A-2

Query Match 82.2%; Score 14.8; DB 3; Length 4403765;
Best Local Similarity 88.9%; Pred. No. 1.6e+02;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GAGAACAAACAAGATTCCG 18
DB 3556350 GAGAACAAACAAGATTCCG 3556367

RESULT 14
US-09-103-840A-1
; Sequence 1, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 4411529
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; OTHER INFORMATION: H37Rv
US-09-103-840A-1

Query Match 82.2%; Score 14.8; DB 3; Length 4411529;
Best Local Similarity 88.9%; Pred. No. 1.6e+02;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GAGAACAAACAAGATTCCG 18
DB 3561497 GAGAACAAACAAGATTCCG 3561514

RESULT 15
US-09-248-796A-13022/C
; Sequence 13022, Application US/09248796A
; Patent No. 6747717
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICANS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 13022
; LENGTH: 204
; TYPE: DNA
; ORGANISM: Candida albicans
US-09-248-796A-13022

Query Match 80.0%; Score 14.4; DB 4; Length 204;
Best Local Similarity 93.8%; Pred. No. 2.7e+02;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 AGAACAAACAAGATTCCG 17
DB 181 AGAACAAACAAGATTCCG 166

RESULT 16
US-09-252-991A-2932
; Sequence 2932, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 2932
; LENGTH: 591
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-2932

Query Match 80.0%; Score 14.4; DB 4; Length 591;
Best Local Similarity 93.8%; Pred. No. 3e+02;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 GAGAACAAACAAGATTCCG 18
DB 11 GAGAACAAACAAGATTCCG 26

RESULT 17
US-09-252-991A-2843
; Sequence 2843, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 2843

Mon Dec 6 12:24:34 2004

us-10-050-189a-6.rn1

Page 13

LENGTH: 684
TYPE: DNA
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-2843

Query Match
Best Local Similarity 93.8%; Score 14.4; DB 4; Length 684;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 GAACAACAAGATTCCG 18
DB 186 GAACAACAAGATTCCG 201

RESULT 18
US-09-270-767-13143/C
Sequence 13143, Application US/09270767
Patent No. 6703491
GENERAL INFORMATION:

APPLICANT: Homburger et al.
TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
FILE REFERENCE: File Reference: 7326-094
CURRENT APPLICATION NUMBER: US/09/270.767
CURRENT FILING DATE: 1999-03-17
NUMBER OF SEQ ID NOS: 62517
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 13143
LENGTH: 771
TYPE: DNA
ORGANISM: Drosophila melanogaster
US-09-270-767-13143

Query Match
Best Local Similarity 93.8%; Score 14.4; DB 4; Length 771;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 AGACAACAAGATTCCG 17
DB 496 AGACAACAAGATTCCG 481

RESULT 19
US-09-577-934A-1
Sequence 1, Application US/09577934A
Patent No. 6590087
GENERAL INFORMATION:

APPLICANT: Bishai, William
TITLE OF INVENTION: WMD, AN ESSENTIAL CELL DIVISION GENE FROM MYCOBACTERIA
FILE REFERENCE: Bishai
CURRENT APPLICATION NUMBER: US/09/577.934A
CURRENT FILING DATE: 2000-05-25
NUMBER OF SEQ ID NOS: 11
SOFTWARE: PatentIn version 3.0
SEQ ID NO 1
LENGTH: 783
TYPE: DNA
ORGANISM: Mycobacterium smegmatis
FEATURE:
NAME/KEY: CDS
LOCATION: (188)..(577)
US-09-577-934A-1

Query Match
Best Local Similarity 93.8%; Score 14.4; DB 4; Length 783;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 GAACAACAAGATTCCG 18
DB 753 GAACAACAAGATTCCG 768

RESULT 20

US-09-252-991A-3123/C
Sequence 3123, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252.991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074.788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094.190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 3123
LENGTH: 867
TYPE: DNA
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-3123

Query Match
Best Local Similarity 93.8%; Score 14.4; DB 4; Length 867;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 GAACAACAAGATTCCG 18
DB 637 GAACAACAAGATTCCG 622

RESULT 21
US-08-956-171E-771/C
Sequence 771, Application US/08956171E
Patent No. 6593114
GENERAL INFORMATION:

APPLICANT: Charles Kunsch
Gil H. Choi
Patrick S. Dillon
Craig A. Rosen
Steven C. Barash
Michael R. Fannon
TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
NUMBER OF SEQUENCES: 5256
CORRESPONDENCE ADDRESS:
ADDRESSER: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4MB storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/956.171E
FILING DATE: 20-Oct-1997
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/009.861
FILING DATE: January 5, 1996
APPLICATION NUMBER: 08/781.986
FILING DATE: January 3, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Mark J. Hyman
REGISTRATION/DOCKET NUMBER: 46,789
TELECOMMUNICATION INFORMATION:
TELEPHONE: (240) 314-1224
TELEFAX: (301) 309-8439
INFORMATION FOR SEQ ID NO: 771:
SEQUENCE CHARACTERISTICS:

LENGTH: 1462 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 771:
US-08-956-171E-771

Query Match 80.0%; Score 14.4; DB 4; Length 1462;
Best Local Similarity 93.8%; Pred. No. 3.4e+02;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 AGAACACACAGATTG 17
DB 816 AGAACACACAGATTG 801

RESULT 22
US-08-781-986A-771/C

Sequence 771, Application US/08781986A

Patent No. 6737248

GENERAL INFORMATION:

APPLICANT: Charles Kunsch

TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences

NUMBER OF SEQUENCES: 5255

CORRESPONDENCE ADDRESS:

ADDRESSEE: Human Genome Sciences, Inc.

STREET: 9410 Key West Avenue

CITY: Rockville

STATE: Maryland

COUNTRY: USA

ZIP: 20850

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage

COMPUTER: HP Vectra 486/33

OPERATING SYSTEM: MSDOS version 6.2

SOFTWARE: ASCII Text

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/781,986A

FILING DATE:

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Benson, Bob

REGISTRATION NUMBER: 30,446

REFERENCE/DOCKET NUMBER: PB248P

TELECOMMUNICATION INFORMATION:

TELEPHONE: (301) 309-8504

TELEFAX: (301) 309-8512

INFORMATION FOR SEQ ID NO: 771:

SEQUENCE CHARACTERISTICS:

LENGTH: 1462 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

US-08-781-986A-771

Query Match 80.0%; Score 14.4; DB 4; Length 1462;
Best Local Similarity 93.8%; Pred. No. 3.4e+02;

Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 AGAACACACAGATTG 17
DB 816 AGAACACACAGATTG 801

RESULT 23

US-09-620-312D-516

Sequence 516, Application US/09620312D

Patent No. 6569662

GENERAL INFORMATION:

APPLICANT: Tang, Y. Tom

APPLICANT: Liu, Chenghua

APPLICANT: Asundi, Vinod

APPLICANT: Zhang, Jie

APPLICANT: Ren, Feiyan

APPLICANT: Chen, Rui-hong

APPLICANT: Zhao, Qing A.

APPLICANT: Weinman, Tom

APPLICANT: Xue, Aidong J.

APPLICANT: Yang, Yonghong

APPLICANT: Wang, Jian-Rui

APPLICANT: Zhou, Ping

APPLICANT: Ma, Yundong

APPLICANT: Wang, Dunrui

APPLICANT: Wang, Zhwei

APPLICANT: John Tillinghast

APPLICANT: Drmanac, Radoje T.

TITLE OF INVENTION: No. 6569662e1 Nucleic Acids and

FILE REFERENCE: 784C1P2B

CURRENT APPLICATION NUMBER: US/09/620,312D

CURRENT FILING DATE: 2000-07-19

PRIOR APPLICATION NUMBER: 09/552,317

PRIOR FILING DATE: 2000-04-25

PRIOR APPLICATION NUMBER: 09/488,725

PRIOR FILING DATE: 2000-01-21

NUMBER OF SEQ ID NOS: 1105

SOFTWARE: pt_FL_genes Version 1.0

SEQ ID NO 516

LENGTH: 1823

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: CDS

LOCATION: (114)..(800)

US-09-620-312D-516

Query Match 80.0%; Score 14.4; DB 4; Length 1823;
Best Local Similarity 93.8%; Pred. No. 3.5e+02;

Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 AGAACACACAGATTG 17
DB 547 AGAACACACAGATTG 562

RESULT 24

US-09-556-601-25

Sequence 25, Application US/09556601

Patent No. 6381599

GENERAL INFORMATION:

APPLICANT: Jefeferies, Thomas W.

APPLICANT: Shi, Nian Qing

TITLE OF INVENTION: SHAW-SENSITIVE TERMINAL OXIDASE GENE FROM

FILE REFERENCE: 96429.9074

CURRENT APPLICATION NUMBER: US/09/556,601

CURRENT FILING DATE: 2000-04-21

NUMBER OF SEQ ID NOS: 29

SOFTWARE: word 97 (DOS text format)

SEQ ID NO 25

LENGTH: 2853

TYPE: DNA

ORGANISM: Pichia stipitis

FEATURE:

NAME/KEY: CDS

LOCATION: (954)..(2027)

US-09-556-601-25

Query Match 80.0%; Score 14.4; DB 3; Length 2853;
Best Local Similarity 93.8%; Pred. No. 3.7e+02;

Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GAGACACACAGATTG 16

Db 1767 GAGAACATCAAGATTTC 1782

RESULT 25
US-09-907-794A-310

; Sequence 310, Application US/09907794A

; Patent No. 6635468

; GENERAL INFORMATION:

; APPLICANT: Genentech, Inc.

; APPLICANT: Ashkenazi, Avi

; APPLICANT: Botstein, David

; APPLICANT: Desnoyers, Luc

; APPLICANT: Eaton, Dan L.

; APPLICANT: Ferrara, Napoleone

; APPLICANT: Filvaroff, Ellen

; APPLICANT: Fong, Sherman

; APPLICANT: Gao, Wei-Qiang

; APPLICANT: Gerber, Hanspeter

; APPLICANT: Gerritsen, Mary E.

; APPLICANT: Goddard, A.

; APPLICANT: Godowski, Paul J.

; APPLICANT: Grimaldi, Christopher J.

; APPLICANT: Gurney, Austin L.

; APPLICANT: Hillan, Kenneth J.

; APPLICANT: Kijavlin, Ivar J.

; APPLICANT: Mather, Jennie P.

; APPLICANT: Paoni, James

; APPLICANT: Paoni, Nicholas F.

; APPLICANT: Roy, Margaret Ann

; APPLICANT: Stewart, Timothy A.

; APPLICANT: Tumas, Daniel

; APPLICANT: Williams, P. Mickey

; APPLICANT: Wood, William, I.

; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic

; FILE REFERENCE: 10466-14

; CURRENT APPLICATION NUMBER: US/09/907,794A

; PRIOR FILING DATE: 2001-07-17

; PRIOR APPLICATION NUMBER: PCT/US00/04414

; PRIOR FILING DATE: 2000-02-22

; PRIOR APPLICATION NUMBER: US 60/143,048

; PRIOR FILING DATE: 1999-07-07

; PRIOR APPLICATION NUMBER: US 60/145,698

; PRIOR FILING DATE: 1999-07-26

; PRIOR APPLICATION NUMBER: US 60/146,222

; PRIOR FILING DATE: 1999-07-28

; PRIOR APPLICATION NUMBER: PCT/US99/20594

; PRIOR FILING DATE: 1999-09-08

; PRIOR APPLICATION NUMBER: PCT/US99/20944

; PRIOR FILING DATE: 1999-09-13

; PRIOR APPLICATION NUMBER: PCT/US99/21090

; PRIOR FILING DATE: 1999-09-15

; PRIOR APPLICATION NUMBER: PCT/US99/21547

; PRIOR FILING DATE: 1999-09-15

; PRIOR APPLICATION NUMBER: PCT/US99/23089

; PRIOR FILING DATE: 1999-10-05

; PRIOR APPLICATION NUMBER: PCT/US99/28214

; PRIOR FILING DATE: 1999-11-29

; PRIOR APPLICATION NUMBER: PCT/US99/28313

; PRIOR FILING DATE: 1999-11-30

; PRIOR APPLICATION NUMBER: PCT/US99/28564

; PRIOR FILING DATE: 1999-12-02

; PRIOR APPLICATION NUMBER: PCT/US99/28565

; PRIOR FILING DATE: 1999-12-02

; PRIOR APPLICATION NUMBER: PCT/US99/30095

; PRIOR FILING DATE: 1999-12-16

; PRIOR APPLICATION NUMBER: PCT/US99/30911

; PRIOR FILING DATE: 1999-12-20

; PRIOR APPLICATION NUMBER: PCT/US99/30999

; PRIOR FILING DATE: 1999-12-20

; PRIOR APPLICATION NUMBER: PCT/US00/00219

; PRIOR FILING DATE: 2000-01-05

; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 310
; LENGTH: 3296
; TYPE: DNA
; ORGANISM: Homo Sapien
US-09-907-794A-310Query Match 80.0%; Score 14.4; DB 4; Length 3296;
Best Local Similarity 93.8%; Pred. No. 3.7e+02;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;CY 2 AGAACATCAAGATTTC 17
Db 2017 AGAACATCAAGATTAG 2032RESULT 26
US-09-905-125A-310

; Sequence 310, Application US/09905125A

; Patent No. 6664376

; GENERAL INFORMATION:

; APPLICANT: Genentech, Inc.

; APPLICANT: Ashkenazi, Avi

; APPLICANT: Botstein, David

; APPLICANT: Desnoyers, Luc

; APPLICANT: Eaton, Dan L.

; APPLICANT: Ferrara, Napoleone

; APPLICANT: Filvaroff, Ellen

; APPLICANT: Fong, Sherman

; APPLICANT: Gao, Wei-Qiang

; APPLICANT: Gerber, Hanspeter

; APPLICANT: Gerritsen, Mary E.

; APPLICANT: Goddard, A.

; APPLICANT: Godowski, Paul J.

; APPLICANT: Grimaldi, Christopher J.

; APPLICANT: Gurney, Austin L.

; APPLICANT: Hillan, Kenneth J.

; APPLICANT: Kijavlin, Ivar J.

; APPLICANT: Mather, Jennie P.

; APPLICANT: Paoni, James

; APPLICANT: Paoni, Nicholas F.

; APPLICANT: Roy, Margaret Ann

; APPLICANT: Stewart, Timothy A.

; APPLICANT: Tumas, Daniel

; APPLICANT: Williams, P. Mickey

; APPLICANT: Wood, William, I.

; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic

; FILE REFERENCE: 10466-14

; CURRENT APPLICATION NUMBER: US/09/905,125A

; PRIOR FILING DATE: 2001-07-12

; PRIOR APPLICATION NUMBER: PCT/US00/04414

; PRIOR FILING DATE: 2000-02-22

; PRIOR APPLICATION NUMBER: US 60/143,048

; PRIOR FILING DATE: 1999-07-07

; PRIOR APPLICATION NUMBER: US 60/145,698

; PRIOR FILING DATE: 1999-07-26

; PRIOR APPLICATION NUMBER: US 60/146,222

; PRIOR FILING DATE: 1999-07-28

; PRIOR APPLICATION NUMBER: PCT/US99/20594

; PRIOR FILING DATE: 1999-09-08

; PRIOR APPLICATION NUMBER: PCT/US99/20944

; PRIOR FILING DATE: 1999-09-13

; PRIOR APPLICATION NUMBER: PCT/US99/21090

; PRIOR FILING DATE: 1999-09-15

; PRIOR APPLICATION NUMBER: PCT/US99/21547

; PRIOR FILING DATE: 1999-09-15

; PRIOR APPLICATION NUMBER: PCT/US99/23089

; PRIOR FILING DATE: 1999-10-05

; PRIOR APPLICATION NUMBER: PCT/US99/28214

; PRIOR FILING DATE: 1999-11-29

; PRIOR APPLICATION NUMBER: PCT/US99/28313

; PRIOR FILING DATE: 1999-11-30


```

; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 310
; LENGTH: 3296
; TYPE: DNA
; ORGANISM: Homo Sapien
; US-09-905-125A-310

Query Match      80.0%; Score 14.4; DB 4; Length 3296;
Best Local Similarity 93.8%; Pred. No. 3.7e+02;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      2 AGACACACAGATTG 17
Db      2017 AGACACACAGATTAG 2032

RESULT 27
US-09-902-775A-310
; Sequence 310, Application US/09902775A
; Patent No. 6686451
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnovers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: 10466-14
; CURRENT APPLICATION NUMBER: US/09/902,775A
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 60/143,048
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
```

```

; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 310
; LENGTH: 3296
; TYPE: DNA
; ORGANISM: Homo Sapien
; US-09-902-775A-310

Query Match      80.0%; Score 14.4; DB 4; Length 3296;
Best Local Similarity 93.8%; Pred. No. 3.7e+02;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      2 AGACACACAGATTG 17
Db      2017 AGACACACAGATTAG 2032

RESULT 28
US-09-906-700-310
; Sequence 310, Application US/09906700
; Patent No. 6723535
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnovers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: 10466-14
```

CURRENT APPLICATION NUMBER: US/09/906,700
PRIOR FILING DATE: 2000-09-18
PRIOR APPLICATION NUMBER: PCT/US00/04414
PRIOR FILING DATE: 2000-02-22
PRIOR APPLICATION NUMBER: US 60/143,048
PRIOR FILING DATE: 1999-07-07
PRIOR APPLICATION NUMBER: US 60/145,698
PRIOR FILING DATE: 1999-07-26
PRIOR APPLICATION NUMBER: US 60/146,222
PRIOR FILING DATE: 1999-07-28
PRIOR APPLICATION NUMBER: PCT/US99/20594
PRIOR FILING DATE: 1999-09-08
PRIOR APPLICATION NUMBER: PCT/US99/20944
PRIOR FILING DATE: 1999-09-13
PRIOR APPLICATION NUMBER: PCT/US99/21090
PRIOR FILING DATE: 1999-09-15
PRIOR APPLICATION NUMBER: PCT/US99/21547
PRIOR FILING DATE: 1999-09-15
PRIOR APPLICATION NUMBER: PCT/US99/23089
PRIOR FILING DATE: 1999-10-05
PRIOR APPLICATION NUMBER: PCT/US99/28214
PRIOR FILING DATE: 1999-11-29
PRIOR APPLICATION NUMBER: PCT/US99/28313
PRIOR FILING DATE: 1999-11-30
PRIOR APPLICATION NUMBER: PCT/US99/28564
PRIOR FILING DATE: 1999-12-02
PRIOR APPLICATION NUMBER: PCT/US99/28565
PRIOR FILING DATE: 1999-12-02
PRIOR APPLICATION NUMBER: PCT/US99/30095
PRIOR FILING DATE: 1999-12-16
PRIOR APPLICATION NUMBER: PCT/US99/30911
PRIOR FILING DATE: 1999-12-20
PRIOR APPLICATION NUMBER: PCT/US99/30999
PRIOR FILING DATE: 1999-12-20
PRIOR APPLICATION NUMBER: PCT/US00/00219
PRIOR FILING DATE: 2000-01-05
NUMBER OF SEQ ID NOS: 423
SEQ ID NO 310
LENGTH: 3296
TYPE: DNA
ORGANISM: Homo Sapien
US-09-906-700-310

Query Match 80.0%; Score 14.4; DB 4;
Best Local Similarity 93.8%; Pred. No. 3.7e+02;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 AGAACACAGATTG 17
Db 2017 AGAACACAGATTG 2032

RESULT 29
US-10-140-002-369
Sequence 369, Application US/10140002
Patent No. 6725730
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Beresini, Naureen
APPLICANT: Defoyse, Laura
APPLICANT: Desnoyers, Luc
APPLICANT: Filvaroff, Ellen
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Sherwood, Steven
APPLICANT: Smith, Victoria
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K
APPLICANT: Wood, William

APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P330R1C59
CURRENT APPLICATION NUMBER: US/10/140,002
PRIOR FILING DATE: 2002-05-06
PRIOR APPLICATION NUMBER: PCT/US99/21090
PRIOR FILING DATE: 1999-09-15
PRIOR APPLICATION NUMBER: PCT/US99/21547
PRIOR FILING DATE: 1999-09-15
PRIOR APPLICATION NUMBER: PCT/US99/23089
PRIOR FILING DATE: 1999-10-05
PRIOR APPLICATION NUMBER: PCT/US99/28214
PRIOR FILING DATE: 1999-11-29
PRIOR APPLICATION NUMBER: PCT/US99/28313
PRIOR FILING DATE: 1999-11-30
PRIOR APPLICATION NUMBER: PCT/US99/28564
PRIOR FILING DATE: 1999-12-02
PRIOR APPLICATION NUMBER: PCT/US99/28565
PRIOR FILING DATE: 1999-12-02
PRIOR APPLICATION NUMBER: PCT/US99/30095
PRIOR FILING DATE: 1999-12-16
PRIOR APPLICATION NUMBER: PCT/US99/30911
PRIOR FILING DATE: 1999-12-20
PRIOR APPLICATION NUMBER: PCT/US99/30999
PRIOR FILING DATE: 1999-12-20
PRIOR APPLICATION NUMBER: PCT/US00/00219
PRIOR FILING DATE: 2000-01-05
NUMBER OF SEQ ID NOS: 550
SEQ ID NO 369
LENGTH: 3296
TYPE: DNA
ORGANISM: Homo Sapien
US-10-140-002-369

QY 2 AGAACACAGATTG 17
Db 2017 AGAACACAGATTG 2032

RESULT 30
US-09-903-603A-310
Sequence 310, Application US/09903603A
Patent No. 676795
GENERAL INFORMATION:

APPLICANT: Genentech, Inc.
APPLICANT: Ashkenazi, Avi
APPLICANT: Botstein, David
APPLICANT: Desnoyers, Luc
APPLICANT: Eaton, Dan L.
APPLICANT: Ferrara, Napoleone
APPLICANT: Filvaroff, Ellen
APPLICANT: Fong, Sherman
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerber, Hanspeter
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, A.
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Hillan, Kenneth, J.
APPLICANT: Kljavin, Ivar J.
APPLICANT: Mather, Jennie P.
APPLICANT: Pan, James
APPLICANT: Paoni, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William, I.
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
FILE REFERENCE: GNE 1618P2C12
CURRENT APPLICATION NUMBER: US/09/903,603A
CURRENT FILING DATE: 2001-07-11
PRIOR APPLICATION NUMBER: PCT/US00/04414
PRIOR FILING DATE: 2000-02-22
PRIOR APPLICATION NUMBER: US 60/143,048
PRIOR FILING DATE: 1999-07-07
PRIOR APPLICATION NUMBER: US 60/145,698
PRIOR FILING DATE: 1999-07-26
PRIOR APPLICATION NUMBER: US 60/146,222
PRIOR FILING DATE: 1999-07-28
PRIOR APPLICATION NUMBER: PCT/US99/20594
PRIOR FILING DATE: 1999-09-08
PRIOR APPLICATION NUMBER: PCT/US99/20944
PRIOR FILING DATE: 1999-09-13
PRIOR APPLICATION NUMBER: PCT/US99/21090
PRIOR FILING DATE: 1999-09-15
PRIOR APPLICATION NUMBER: PCT/US99/21547

PRIOR FILING DATE: 1999-09-15
PRIOR APPLICATION NUMBER: PCT/US99/23089
PRIOR FILING DATE: 1999-10-05 PCT/US99/28214
PRIOR APPLICATION NUMBER: PCT/US99/28214
PRIOR FILING DATE: 1999-11-29 PCT/US99/28313
PRIOR APPLICATION NUMBER: PCT/US99/28313
PRIOR FILING DATE: 1999-11-30 PCT/US99/28564
PRIOR APPLICATION NUMBER: PCT/US99/28564
PRIOR FILING DATE: 1999-12-02 PCT/US99/28565
PRIOR APPLICATION NUMBER: PCT/US99/28565
PRIOR FILING DATE: 1999-12-02 PCT/US99/30095
PRIOR APPLICATION NUMBER: PCT/US99/30095
PRIOR FILING DATE: 1999-12-16 PCT/US99/30911
PRIOR APPLICATION NUMBER: PCT/US99/30911
PRIOR FILING DATE: 1999-12-20 PCT/US99/30999
PRIOR APPLICATION NUMBER: PCT/US99/30999
PRIOR FILING DATE: 1999-12-20 PCT/US00/00219
PRIOR APPLICATION NUMBER: PCT/US00/00219
NUMBER OF SEQ ID NOS: 423
SEQ ID NO 310
LENGTH: 3296
TYPE: DNA
ORGANISM: Homo Sapien
US-09-903-603A-310

Query Match 80.0%; Score 14.4; DB 4; Length 3296;
Best Local Similarity 93.8%; Pred. No. 3.7e+02;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 AGAACACACAGATTG 17
DB 2017 AGAACACACAGATTG 2032

RESULT 31
US-08-936-171E-228
Sequence 228, Application US/08956171E
Patent No. 6593114
GENERAL INFORMATION:

APPLICANT: Charles Kunesh

Gill H. Choi

Patrick S. Dillon

Craig A. Rosen

Steven C. Barash

Michael R. Fannon

ADDRESSER: Human Genome Sciences, Inc.

STREET: 9410 Key West Avenue

CITY: Rockville

STATE: Maryland

COUNTRY: USA

ZIP: 20850

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.50 inch, 1.4MB storage

COMPUTER: HP Vectra 486/33

OPERATING SYSTEM: MSDOS version 6.2

SOFTWARE: ASCII Text

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/956,171E

FILING DATE: 20-Oct-1997

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/009,861
FILING DATE: January 5, 1996
APPLICATION NUMBER: 08/781,986
FILING DATE: January 3, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Mark J. Hyman
REGISTRATION NUMBER: 46,789
REFERENCE/DOCKET NUMBER: PB248P1

TELECOMMUNICATION INFORMATION:
TELEPHONE: (240) 314-1224
TELEFAX: (301) 309-8439
INFORMATION FOR SEQ ID NO: 228:
SEQUENCE CHARACTERISTICS:
LENGTH: 4488 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 228:
US-08-956-171E-228

Query Match 80.0%; Score 14.4; DB 4; Length 4488;
Best Local Similarity 93.8%; Pred. No. 3.9e+02;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 GAGAACACACAGATTG 16
DB 111 GAGAACACACAGATTG 126

RESULT 32
US-08-781-986A-228
Sequence 228, Application US/08781-986A
Patent No. 6737248
GENERAL INFORMATION:

APPLICANT: Charles Kunesh

TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences

NUMBER OF SEQUENCES: 5255

CORRESPONDENCE ADDRESS:

ADDRESSER: Human Genome Sciences, Inc.

STREET: 9410 Key West Avenue

CITY: Rockville

STATE: Maryland

COUNTRY: USA

ZIP: 20850

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.50 inch, 1.4MB storage

COMPUTER: HP Vectra 486/33

OPERATING SYSTEM: MSDOS version 6.2

SOFTWARE: ASCII Text

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/781,986A

FILING DATE:

CLASSIFICATION: 435

PRIOR APPLICATION NUMBER:

ATTORNEY/AGENT INFORMATION:

NAME: Benson, Bob

REGISTRATION NUMBER: 30,446

REFERENCE/DOCKET NUMBER: PB248BP

TELECOMMUNICATION INFORMATION:

TELEPHONE: (301) 309-8504

TELEFAX: (301) 309-8512

INFORMATION FOR SEQ ID NO: 228:
SEQUENCE CHARACTERISTICS:
LENGTH: 4488 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear

US-08-781-986A-228

Query Match 80.0%; Score 14.4; DB 4; Length 4488;
Best Local Similarity 93.8%; Pred. No. 3.9e+02;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 GAGAACACACAGATTG 16
DB 111 GAGAACACACAGATTG 126

RESULT 33

```
US-09-345-882-1
Sequence 1, Application US/09345882
Patent No. 639373
GENERAL INFORMATION:
APPLICANT: Bouquelere, Lydie
TITLE OF INVENTION: A NUCLEIC ACID ENCODING A RETINOBLASTOMA BINDING PROTEIN (RBP-7)
FILE REFERENCE: GENSET.031A
CURRENT APPLICATION NUMBER: US/09/345, 882
CURRENT FILING DATE: 1999-06-30
PRIOR APPLICATION NUMBER: US 60/091,315
PRIOR FILING DATE: 1998-06-30
PRIOR APPLICATION NUMBER: US 60/111,909
PRIOR FILING DATE: 1998-12-10
NUMBER OF SEQ ID NOS: 140
SOFTWARE: Patent.pm
SEQ ID NO 1
LENGTH: 162450
TYPE: DNA
ORGANISM: Homo sapiens
FEATURES:
NAME/KEY: allele
LOCATION: 72794
OTHER INFORMATION: 5-124-273 : polymorphic base A or G
FEATURE:
NAME/KEY: allele
LOCATION: 88073
OTHER INFORMATION: 5-127-261 : polymorphic base A or C
FEATURE:
NAME/KEY: allele
LOCATION: 90842
OTHER INFORMATION: 99-1437-325 : polymorphic base A or G
FEATURE:
NAME/KEY: allele
LOCATION: 93714
OTHER INFORMATION: 5-128-60 : polymorphic base deletion of GT
FEATURE:
NAME/KEY: allele
LOCATION: 97122
OTHER INFORMATION: 99-1442-224 : polymorphic base G or T
FEATURE:
NAME/KEY: allele
LOCATION: 97152
OTHER INFORMATION: 5-129-144 : polymorphic base deletion of T
FEATURE:
NAME/KEY: allele
LOCATION: 99098
OTHER INFORMATION: 5-130-257 : polymorphic base A or G
FEATURE:
NAME/KEY: allele
LOCATION: 99117
OTHER INFORMATION: 5-130-276 : polymorphic base A or G
FEATURE:
NAME/KEY: allele
LOCATION: 103806
OTHER INFORMATION: 5-131-395 : polymorphic base A or T
FEATURE:
NAME/KEY: allele
LOCATION: 106940
OTHER INFORMATION: 5-133-375 : polymorphic base insertion of A
FEATURE:
NAME/KEY: allele
LOCATION: 108106
OTHER INFORMATION: 5-135-155 : polymorphic base insertion of A
FEATURE:
NAME/KEY: allele
LOCATION: 108149
OTHER INFORMATION: 5-135-198 : polymorphic base insertion of GTTT
FEATURE:
NAME/KEY: allele
LOCATION: 108308
OTHER INFORMATION: 5-135-357 : polymorphic base A or G
FEATURE:
NAME/KEY: allele
LOCATION: 108471
OTHER INFORMATION: 5-136-174 : polymorphic base C or T
FEATURE:
NAME/KEY: allele
LOCATION: 134134
OTHER INFORMATION: 5-140-120 : polymorphic base C or T
FEATURE:
NAME/KEY: allele
LOCATION: 134362
OTHER INFORMATION: 5-140-348 : polymorphic base insertion of A
FEATURE:
NAME/KEY: allele
LOCATION: 134374
OTHER INFORMATION: 5-140-361 : polymorphic base insertion of CA
FEATURE:
NAME/KEY: allele
LOCATION: 146328
OTHER INFORMATION: 5-143-84 : polymorphic base A or G
FEATURE:
NAME/KEY: allele
LOCATION: 146345
OTHER INFORMATION: 5-143-101 : polymorphic base A or C
FEATURE:
NAME/KEY: allele
LOCATION: 150329
OTHER INFORMATION: 5-145-24 : polymorphic base A or G
FEATURE:
NAME/KEY: allele
LOCATION: 160031
OTHER INFORMATION: 5-148-352 : polymorphic base G or T
FEATURE:
NAME/KEY: allele
LOCATION: 72771..72817
OTHER INFORMATION: polymorphic fragment 5-124-273 SEQ ID51
FEATURE:
NAME/KEY: allele
LOCATION: 88050..88096
OTHER INFORMATION: polymorphic fragment 5-127-261 SEQ ID52
FEATURE:
NAME/KEY: allele
LOCATION: 90819..90865
OTHER INFORMATION: complement polymorphic fragment 99-1437-325 SEQ ID49
FEATURE:
NAME/KEY: allele
LOCATION: 93690..93736
OTHER INFORMATION: polymorphic fragment 5-128-60 SEQ ID32
FEATURE:
NAME/KEY: allele
LOCATION: 93690..93736
OTHER INFORMATION: polymorphic fragment 5-128-60 SEQ ID53
FEATURE:
NAME/KEY: allele
LOCATION: 97099..97145
OTHER INFORMATION: polymorphic fragment 99-1442-224 SEQ ID50
FEATURE:
NAME/KEY: allele
LOCATION: 97099..97145
OTHER INFORMATION: polymorphic fragment 99-1442-224 SEQ ID71
FEATURE:
NAME/KEY: allele
```

```
LOCATION: 97130..97177
OTHER INFORMATION: polymorphic fragment 5-129-144 SEQ ID33
FEATURE:
NAME/KEY: allele
LOCATION: 97130..97177
OTHER INFORMATION: polymorphic fragment 5-129-144 SEQ ID34
FEATURE:
NAME/KEY: allele
LOCATION: 99075..99121
OTHER INFORMATION: polymorphic fragment 5-130-257 SEQ ID34
FEATURE:
NAME/KEY: allele
LOCATION: 99094..99140
OTHER INFORMATION: polymorphic fragment 5-130-276 SEQ ID35
FEATURE:
NAME/KEY: allele
LOCATION: 99075..99121
OTHER INFORMATION: polymorphic fragment 5-130-257 SEQ ID55
FEATURE:
NAME/KEY: allele
LOCATION: 99094..99140
OTHER INFORMATION: polymorphic fragment 5-130-276 SEQ ID56
FEATURE:
NAME/KEY: allele
LOCATION: 103783..103828
OTHER INFORMATION: polymorphic fragment 5-131-395 SEQ ID36
FEATURE:
NAME/KEY: allele
LOCATION: 103783..103828
OTHER INFORMATION: polymorphic fragment 5-131-395 SEQ ID57
FEATURE:
NAME/KEY: allele
LOCATION: 106918..106966
OTHER INFORMATION: polymorphic fragment 5-133-375 SEQ ID37
FEATURE:
NAME/KEY: allele
LOCATION: 106918..106966
OTHER INFORMATION: polymorphic fragment 5-133-375 SEQ ID58
FEATURE:
NAME/KEY: allele
LOCATION: 108084..108130
OTHER INFORMATION: polymorphic fragment 5-135-155 SEQ ID59
FEATURE:
NAME/KEY: allele
LOCATION: 108127..108177
OTHER INFORMATION: polymorphic fragment 5-135-198 SEQ ID39
FEATURE:
NAME/KEY: allele
LOCATION: 108127..108177
OTHER INFORMATION: polymorphic fragment 5-135-198 SEQ ID60
FEATURE:
Query Match 80.0%; Score 14.4; DB 3; Length 162450;
Best Local Similarity 93.8%; Pred. No. 5.6e+02;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GAGAACAAACAAGATTC 16
Db 95472 GACAAACAAGATTC 95487

RESULT 34
US-09-103-840A-2/c
Sequence 2, Application US/09103840A
Patent No. 6294328
GENERAL INFORMATION:
APPLICANT: FLEISCHMAN, Robert D.
APPLICANT: WHITE, Owen R.
APPLICANT: FRASER, Claire M.
```

```
APPLICANT: VENTER, John C.
TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
TITLE OF INVENTION: TUBERCULOSIS
FILE REFERENCE: 24366-20007.00
CURRENT APPLICATION NUMBER: US/09/103,840A
CURRENT FILING DATE: 1998-06-24
NUMBER OF SEQ ID NOS: 2
SOFTWARE: Patent In Ver. 2.1
SEQ ID NO 2
LENGTH: 4403765
TYPE: DNA
ORGANISM: Mycobacterium tuberculosis
FEATURE:
OTHER INFORMATION: CDC 1551
OTHER INFORMATION: "t" bases at various positions throughout the sequence
OTHER INFORMATION: represent a, t, c or g
US-09-103-840A-2

Query Match 80.0%; Score 14.4; DB 3; Length 4403765;
Best Local Similarity 93.8%; Pred. No. 1.6e+02;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GAGAACAAACAAGATTC 16
Db 2998437 GCGAACAAACAAGATTC 2998422

RESULT 35
US-09-103-840A-1/c
Sequence 1, Application US/09103840A
Patent No. 6294328
GENERAL INFORMATION:
APPLICANT: FLEISCHMAN, Robert D.
APPLICANT: WHITE, Owen R.
APPLICANT: FRASER, Claire M.
APPLICANT: VENTER, John C.
TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
TITLE OF INVENTION: TUBERCULOSIS
FILE REFERENCE: 24366-20007.00
CURRENT APPLICATION NUMBER: US/09/103,840A
CURRENT FILING DATE: 1998-06-24
NUMBER OF SEQ ID NOS: 2
SOFTWARE: Patent In Ver. 2.1
SEQ ID NO 1
LENGTH: 4411529
TYPE: DNA
ORGANISM: Mycobacterium tuberculosis
OTHER INFORMATION: H37Rv
US-09-103-840A-1

Query Match 80.0%; Score 14.4; DB 3; Length 4411529;
Best Local Similarity 93.8%; Pred. No. 1.6e+02;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GAGAACAAACAAGATTC 16
Db 3003531 GCGAACAAACAAGATTC 3003516

RESULT 36
US-09-252-991A-15780
Sequence 15780, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
TITLE OF INVENTION: AERUSINOSA FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
```

;; PRIOR FILING DATE: 1998-07-27
;; NUMBER OF SEQ ID NOS: 33142
;; SEQ ID NO: 15780
;; LENGTH: 723
;; TYPE: DNA
;; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-15780

Query Match 77.8%; Score 14; DB 4; Length 723;
Best Local Similarity 100.0%; Pred. No. 5e+02;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GAGACACACAGAT 14
Db 11 GAGACACACAGAT 24

RESULT 37
US-09-248-796A-662
; Sequence 662, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; PRIOR FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 662
; LENGTH: 792
; TYPE: DNA
; ORGANISM: Candida albicans
US-09-248-796A-662

Query Match 77.8%; Score 14; DB 4; Length 792;
Best Local Similarity 100.0%; Pred. No. 5e+02;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 AGAACACACAGATT 15
Db 777 AGAACACACAGATT 790

RESULT 38
US-09-252-991A-15910/C
; Sequence 15910, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 15910
; LENGTH: 825
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-15910

Query Match 77.8%; Score 14; DB 4; Length 825;
Best Local Similarity 100.0%; Pred. No. 5e+02;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GAGACACACAGAT 14
Db 732 GAGACACACAGAT 719

RESULT 39
US-09-248-796A-10471
; Sequence 10471, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 10471
; LENGTH: 969
; TYPE: DNA
; ORGANISM: Candida albicans
US-09-248-796A-10471

Query Match 77.8%; Score 14; DB 4; Length 969;
Best Local Similarity 100.0%; Pred. No. 5.1e+02;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 AGAACACACAGATT 15
Db 252 AGAACACACAGATT 265

RESULT 40
US-09-252-991A-15881/C
; Sequence 15881, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 15881
; LENGTH: 1083
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-15881

Query Match 77.8%; Score 14; DB 4; Length 1083;
Best Local Similarity 100.0%; Pred. No. 5.2e+02;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GAGACACACAGAT 14
Db 1018 GAGACACACAGAT 1005

RESULT 41
US-09-252-991A-15836
; Sequence 15836, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:

```
/ APPLICANT: Marc J. Rubenfield et al.
/ TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
/ TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
/ FILE REFERENCE: 107196.136
/ CURRENT APPLICATION NUMBER: US/09/252,991A
/ PRIOR FILING DATE: 1999-02-18
/ PRIOR APPLICATION NUMBER: US 60/074,788
/ PRIOR FILING DATE: 1998-02-18
/ PRIOR APPLICATION NUMBER: US 60/094,190
/ PRIOR FILING DATE: 1998-07-27
/ NUMBER OF SEQ ID NOS: 33142
/ SEQ ID NO 15836
/ LENGTH: 1317
/ TYPE: DNA
/ ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-15836

Query Match 77.8%; Score 14; DB 4; Length 1317;
Best Local Similarity 100.0%; Pred. No. 5.3e+02;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAGAACACACAGAT 14
Db 1293 GAGAACACACAGAT 1306

RESULT 42
US-09-252-991A-15811
/ Sequence 15811, Application US/09252991A
/ Patent No. 6551785
/ GENERAL INFORMATION:
/ APPLICANT: Marc J. Rubenfield et al.
/ TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
/ TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
/ FILE REFERENCE: 107196.136
/ CURRENT APPLICATION NUMBER: US/09/252,991A
/ CURRENT FILING DATE: 1999-02-18
/ PRIOR APPLICATION NUMBER: US 60/074,788
/ PRIOR FILING DATE: 1998-02-18
/ PRIOR APPLICATION NUMBER: US 60/094,190
/ PRIOR FILING DATE: 1998-07-27
/ NUMBER OF SEQ ID NOS: 33142
/ SEQ ID NO 15811
/ LENGTH: 2056
/ TYPE: DNA
/ ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-15811

Query Match 77.8%; Score 14; DB 4; Length 2056;
Best Local Similarity 100.0%; Pred. No. 5.6e+02;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAGAACACACAGAT 14
Db 19 GAGAACACACAGAT 32

RESULT 43
US-09-809-665A-39
/ Sequence 39, Application US/09809665A
/ Patent No. 6790950
/ GENERAL INFORMATION:
/ APPLICANT: Lowery E., David, et al.
/ TITLE OF INVENTION: Anti-Bacterial Vaccine Compositions
/ FILE REFERENCE: 28141/00435
/ CURRENT APPLICATION NUMBER: US/09/809,665A
/ CURRENT FILING DATE: 2001-03-15
/ PRIOR APPLICATION NUMBER: 60/153,453
/ PRIOR FILING DATE: 1999-09-10
/ PRIOR APPLICATION NUMBER: 60/128,689
/ PRIOR FILING DATE: 1999-04-09
/ PRIOR APPLICATION NUMBER: 09/545,199
/ PRIOR FILING DATE: 2000-04-06
```

```
/ NUMBER OF SEQ ID NOS: 197
/ SOFTWARE: Patentln Ver. 2.0
/ SEQ ID NO 39
/ LENGTH: 2628
/ TYPE: DNA
/ ORGANISM: Pasteurella multocida
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (326)..(766)
/ OTHER INFORMATION: mloc
US-09-809-665A-39

Query Match 77.8%; Score 14; DB 4; Length 2628;
Best Local Similarity 100.0%; Pred. No. 5.8e+02;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 AGAACACACAGAT 15
Db 406 AGAACACACAGAT 419

RESULT 44
US-08-956-171E-1643
/ Sequence 1643, Application US/08956171E
/ Patent No. 6593114
/ GENERAL INFORMATION:
/ APPLICANT: Charles Kunsch
/ GIL H. Choi
/ Patrick S. Dillon
/ Craig A. Rosen
/ Steven C. Barash
/ Michael R. Farnon
/ TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
/ NUMBER OF SEQUENCES: 5256
/ CORRESPONDENCE ADDRESS:
/ ADDRESSER: Human Genome Sciences, Inc.
/ STREET: 9410 Key West Avenue
/ CITY: Rockville
/ STATE: Maryland
/ COUNTRY: USA
/ ZIP: 20850
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Diskette, 3.50 inch, 1.44Mb storage
/ COMPUTER: HP Vectra 486/33
/ OPERATING SYSTEM: MSDOS version 6.2
/ SOFTWARE: ASCII Text
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/956,171E
/ FILING DATE: 20-Oct-1997
/ CLASSIFICATION: <Unknown>
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 60/009,861
/ FILING DATE: January 5, 1996
/ APPLICATION NUMBER: 08/781,986
/ FILING DATE: January 3, 1997
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Mark J. Hyman
/ REGISTRATION NUMBER: 46,789
/ REFERENCE/DOCKET NUMBER: PB248P1
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (240) 314-1224
/ TELEFAX: (301) 309-8439
/ INFORMATION FOR SEQ ID NO: 1643:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 66 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: double
/ TOPOLOGY: linear
/ SEQUENCE DESCRIPTION: SEQ ID NO: 1643:
US-08-956-171E-1643

Query Match 76.7%; Score 13.8; DB 4; Length 66;
Best Local Similarity 88.2%; Pred. No. 4.7e+02;
```


Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 AGAACACACAGATTCCG 18
| | | | | | | | | | | | | | | | | |
Db 42 AGAACACATAGACATCCG 58

RESULT 45

US-08-781-986A-1643
; Sequence 1643, Application US/08781986A
; Patent No. 6737248
; GENERAL INFORMATION:
; APPLICANT: Charles Kunach
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 5255
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/781,986A
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Benson, Bob
; REGISTRATION NUMBER: 30,446
; REFERENCE/DOCKET NUMBER: PB248PP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 1643:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 66 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
US-08-781-986A-1643

Query Match 76.7%; Score 13.8; DB 4; Length 66;

Best Local Similarity 88.2%; Pred. No. 4.7e+02;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 AGAACACACAGATTCCG 18
| | | | | | | | | | | | | | | | | |
Db 42 AGAACACATAGACATCCG 58

RESULT 46

US-09-222-575-44
; Sequence 44, Application US/09222575
; Patent No. 6387697
; GENERAL INFORMATION:
; APPLICANT: Yugu, Jiang
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jianshun
; TITLE OF INVENTION: Compositions for the Treatment and Diagnosis of Breast Cancer
; TITLE OF INVENTION: and Methods for their Use
; FILE REFERENCE: 210121.470
; CURRENT APPLICATION NUMBER: US/09/222,575
; CURRENT FILING DATE: 1998-12-28
; NUMBER OF SEQ ID NOS: 174

; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 44
; LENGTH: 160
; TYPE: DNA
; ORGANISM: Human
US-09-222-575-44

QY 2 AGAACACACAGATTCCG 18
| | | | | | | | | | | | | | | | | |
Db 45 AGAACACACAGATTCCG 61

Query Match 76.7%; Score 13.8; DB 3; Length 160;
Best Local Similarity 88.2%; Pred. No. 5.2e+02;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 AGAACACACAGATTCCG 18
| | | | | | | | | | | | | | | | | |
Db 45 AGAACACACAGATTCCG 61

RESULT 47

US-09-389-681-44
; Sequence 44, Application US/09389681A
; Patent No. 6518237
; GENERAL INFORMATION:
; APPLICANT: Yugu, Jiang
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jianshun
; TITLE OF INVENTION: COMPOSITIONS FOR THE TREATMENT AND
; TITLE OF INVENTION: DIAGNOSIS OF BREAST CANCER AND METHODS FOR THEIR USE
; FILE REFERENCE: 210121.470C3
; CURRENT APPLICATION NUMBER: US/09/389,681A
; CURRENT FILING DATE: 1999-09-02
; NUMBER OF SEQ ID NOS: 463
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 44
; LENGTH: 160
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-389-681-44

Query Match 76.7%; Score 13.8; DB 4; Length 160;
Best Local Similarity 88.2%; Pred. No. 5.2e+02;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 AGAACACACAGATTCCG 18
| | | | | | | | | | | | | | | | | |
Db 45 AGAACACACAGATTCCG 61

RESULT 48

US-09-620-405B-44
; Sequence 44, Application US/09620405B
; Patent No. 6528054
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yugu
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jianshun
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Hepler, William T.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF BREAST CANCER
; FILE REFERENCE: 210121.470C8
; CURRENT APPLICATION NUMBER: US/09/620,405B
; CURRENT FILING DATE: 2000-07-20
; NUMBER OF SEQ ID NOS: 495
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 44
; LENGTH: 160
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-620-405B-44

Query Match 76.7%; Score 13.8; DB 4; Length 160;
Best Local Similarity 88.2%; Pred. No. 5.2e+02;

Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 AGAACACACAGATTGCG 18
 Db 45 AGAACACACAGATCCAC 61

RESULT 49

US-09-339-338-44
 ; Sequence 44, Application US/09339338A
 ; Patent No. 6573368
 ; GENERAL INFORMATION:
 ; APPLICANT: Yugu, Jiang
 ; APPLICANT: Dillon, Davin C.
 ; APPLICANT: Mitcham, Jennifer L.
 ; APPLICANT: Xu, Jiangchun
 ; TITLE OF INVENTION: COMPOSITIONS FOR THE TREATMENT AND
 ; FILE REFERENCE: 210121.470C2
 ; CURRENT APPLICATION NUMBER: US/09/339,338A
 ; CURRENT FILING DATE: 1999-06-23
 ; NUMBER OF SEQ ID NOS: 315
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 44
 ; LENGTH: 160
 ; TYPE: DNA
 ; ORGANISM: Homo sapien
 US-09-339-338-44

Query Match 76.7%; Score 13.8; DB 4; Length 160;

Best Local Similarity 88.2%; Pred. No. 5.2e+02;
 Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 AGAACACACAGATTGCG 18
 Db 45 AGAACACACAGATCCAC 61

RESULT 50

US-09-433-826B-44
 ; Sequence 44, Application US/09433826B
 ; Patent No. 6579973
 ; GENERAL INFORMATION:
 ; APPLICANT: Jiang, Yugu
 ; APPLICANT: Dillon, Davin C.
 ; APPLICANT: Mitcham, Jennifer L.
 ; APPLICANT: Xu, Jiangchun
 ; APPLICANT: Harlocker, Susan L.
 ; TITLE OF INVENTION: COMPOSITIONS FOR THE TREATMENT AND
 ; FILE REFERENCE: 210121.470C4
 ; CURRENT APPLICATION NUMBER: US/09/433,826B
 ; CURRENT FILING DATE: 1999-11-03
 ; NUMBER OF SEQ ID NOS: 474
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 44
 ; LENGTH: 160
 ; TYPE: DNA
 ; ORGANISM: Homo sapien
 US-09-433-826B-44

Query Match 76.7%; Score 13.8; DB 4; Length 160;

Best Local Similarity 88.2%; Pred. No. 5.2e+02;
 Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 AGAACACACAGATTGCG 18
 Db 45 AGAACACACAGATCCAC 61

Search completed: December 3, 2004, 05:51:27
 Job time : 123.684 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 3, 2004, 01:41:45 ; Search time 2282.45 Seconds
(without alignments)
287.374 Million cell updates/sec

Title: US-10-050-189A-6

Perfect score: 18

Sequence: 1 gagagacacacagatcgc 18

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 32822875 segs, 18219865908 residues

Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 1000 summaries

Database :

EST: *
1: gb_est1: *
2: gb_est2: *
3: gb_hic: *
4: gb_est3: *
5: gb_est4: *
6: gb_est5: *
7: gb_est6: *
8: gb_est7: *
9: gb_est8: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result	No.	Score	Query	Match	Length	ID	Description
	1	17	94.4	164	9	CG665075	OST452977
	2	17	94.4	383	9	CG559318	OST178045
	3	17	94.4	936	9	CNS07414	AL428222 clone BAO
	4	16.4	91.1	179	4	BG410386	BG410386 947045H08
	5	16.4	91.1	357	6	CD460790	F909 01n0
	6	16.4	91.1	481	8	AO210247	HS_3115 B
	7	16.4	91.1	504	6	CD657925	EST659
	8	16.4	91.1	514	4	BM076070	WEST364-D
	9	16.4	91.1	768	7	COB17404	PA_SEA001
	10	16.4	91.1	1138	4	BI823206	603039469
	11	16	88.9	426	9	CE085246	Liqr-g88-
	12	16	88.9	426	9	CE085246	Liqr-g88-
	13	16	88.9	646	7	CK130476	AT08181.3
	14	16	88.9	690	8	BZ178809	CH230-443
	15	16	88.9	744	8	CC082521	CSU-K31T
	16	16	88.9	892	9	BU017470	QHE15P17.
	17	15.4	85.6	229	5	BU017470	QHE15P17.
	18	15.4	85.6	229	5	BU017470	QHE15P17.
	19	15.4	85.6	269	5	BU016050	QHE11P07.
	20	15.4	85.6	270	6	CD944524	RDJ_72 Ge
	21	15.4	85.6	270	6	CD944524	RDJ_72 Ge
	22	15.4	85.6	270	6	CD967114	SES_132 G
	23	15.4	85.6	270	6	CD967248	SES_61 Ge
	24	15.4	85.6	273	2	AW977559	EST389668

25	15.4	85.6	284	5	BO910769	CHAI5C10.
26	15.4	85.6	285	5	BU016309	QHE12117.
27	15.4	85.6	286	5	BU021781	QHE4E07.Y
28	15.4	85.6	300	6	CA485884	WHE4324.C
29	15.4	85.6	304	2	AW646597	PM2-SN001
30	15.4	85.6	312	1	AV216565	AV216565
31	15.4	85.6	340	1	AW787371	n2508a1.x
32	15.4	85.6	351	2	AW447131	88282 YAR
33	15.4	85.6	365	1	AU257195	AU257195
34	15.4	85.6	366	1	BZ651766	OCAMP82TC
35	15.4	85.6	367	2	BE851239	IL5-EN008
36	15.4	85.6	380	3	AO851814	Arm-genes
37	15.4	85.6	383	8	AO851814	LMaJFV1_1
38	15.4	85.6	400	1	AA994392	ou10g03.s
39	15.4	85.6	402	9	CG722358	1119071FO
40	15.4	85.6	405	1	AI280419	q195d08.x
41	15.4	85.6	409	1	AI220277	q573d03.x
42	15.4	85.6	413	1	AA907628	cm10F08.s
43	15.4	85.6	419	4	BM173099	900453 AY
44	15.4	85.6	428	1	AA612092	cb40a06.s
45	15.4	85.6	430	8	BZ769778	SALK_1427
46	15.4	85.6	431	5	BP645308	BP645308
47	15.4	85.6	433	4	BG543327	E0832 Chi
48	15.4	85.6	437	4	BM360109	CA_Pa002
49	15.4	85.6	438	1	AA610098	AI19e07.s
50	15.4	85.6	438	4	BM801478	BM801478
51	15.4	85.6	466	1	AA772688	a135b05.s
52	15.4	85.6	474	1	AI266637	q45b11.x
53	15.4	85.6	478	7	CN261424	170005322
54	15.4	85.6	478	8	BH849056	SAIK_0582
55	15.4	85.6	490	4	BUT59781	BUT59781
56	15.4	85.6	498	8	BH179757	O15_D-23-
57	15.4	85.6	498	9	CNS07LRN	AL616709 T7 end of
58	15.4	85.6	505	4	BJ108381	BJ108381
59	15.4	85.6	512	4	BG103963	RH122_38
60	15.4	85.6	516	4	BUT60375	BUT60375
61	15.4	85.6	517	4	BG103332	RH122_18
62	15.4	85.6	519	8	AO889206	HS_5539_A
63	15.4	85.6	525	7	CF391508	RTDR3_9_E
64	15.4	85.6	537	2	AW301203	xss57905.x
65	15.4	85.6	537	4	BJ765562	BJ765562
66	15.4	85.6	539	4	BJ753442	BJ753442
67	15.4	85.6	543	4	BJ772991	BJ772991
68	15.4	85.6	548	4	BJ126386	BJ126386
69	15.4	85.6	551	4	BG355639	EM1_18 DO
70	15.4	85.6	552	4	BJ112454	BJ112454
71	15.4	85.6	553	6	CD991044	QAZ3610.Y
72	15.4	85.6	555	1	AI202762	q145d06.x
73	15.4	85.6	561	5	BO664424	HVO3B04u
74	15.4	85.6	565	4	BM323691	PTC1_22 B
75	15.4	85.6	565	6	CB015597	HC_d11.42
76	15.4	85.6	566	4	BG465737	RH122_48-
77	15.4	85.6	567	1	AI871351	AI871351
78	15.4	85.6	569	1	AI947649	603026A07
79	15.4	85.6	572	5	BO664177	HVO2D11u
80	15.4	85.6	576	1	AI220156	q579b11.x
81	15.4	85.6	579	4	BI508873	BI508873
82	15.4	85.6	580	4	BM323691	PTC1_22 B
83	15.4	85.6	585	2	AW015551	UI-H-B10P
84	15.4	85.6	587	1	AA775336	ad13604.s
85	15.4	85.6	587	7	CO790032	NT008C DO
86	15.4	85.6	588	2	AW165994	x43e604.x
87	15.4	85.6	588	6	CF022808	CE022808
88	15.4	85.6	588	9	CE702297	Liqr-g88-
89	15.4	85.6	598	2	BJ117751	BJ117751
90	15.4	85.6	600	2	AW146888	614069C12
91	15.4	85.6	603	3	BJ477134	BU477134
92	15.4	85.6	622	3	AY613879	MesenchyC
93	15.4	85.6	624	4	BJ753364	BJ753364
94	15.4	85.6	625	5	BU573976	PA_Ba000
95	15.4	85.6	625	8	BH745756	G271d08.b
96	15.4	85.6	628	4	BU108119	BU108119
97	15.4	85.6	630	5	BP117787	BP117787

98	15.4	85.6	631	4	BG718181	602696145	171	15	83.3	521	1	AV671114
99	15.4	85.6	657	7	CS511600	rswwdc.00	172	15	83.3	521	2	BE601652
100	15.4	85.6	659	4	BJ770602	Bj770602	173	15	83.3	523	6	CD309974
101	15.4	85.6	663	6	CB332921	py70f07.y	174	15	83.3	539	1	AV555802
102	15.4	85.6	664	5	BQ997128	Q6G14M24.	175	15	83.3	545	6	CD823327
103	15.4	85.6	664	7	CF391534	RTDR3_9_E	176	15	83.3	546	5	BQ091272
104	15.4	85.6	674	1	A1356774	qy2h06.x	177	15	83.3	550	6	CB261944
105	15.4	85.6	679	8	BH32217	od122d04.	178	15	83.3	553	5	BM453301
106	15.4	85.6	688	4	BG495794	602539050	179	15	83.3	561	5	BM413225
107	15.4	85.6	717	9	AG456384	Mue muscu	180	15	83.3	562	4	B1397065
108	15.4	85.6	737	8	BZ651772	OGANP82TM	181	15	83.3	573	1	AA558907
109	15.4	85.6	739	5	BT725101	SUNCIH06	182	15	83.3	576	5	BM494170
110	15.4	85.6	746	9	CL667846	PR10156a_	183	15	83.3	593	1	AV561026
111	15.4	85.6	754	9	AG309135	Mue muscu	184	15	83.3	593	3	BM114659
112	15.4	85.6	762	6	CB954477	AGENCOURT	185	15	83.3	594	1	AJ8033959
113	15.4	85.6	764	9	CC704366	OGWE14TH	186	15	83.3	599	1	BM429329
114	15.4	85.6	773	7	CO484316	G00205.B3	187	15	83.3	601	1	AV888715
115	15.4	85.6	775	3	AY223366	Schistoso	188	15	83.3	601	6	CA555346
116	15.4	85.6	777	8	BZ479818	BONCDSTR	189	15	83.3	602	6	CB371276
117	15.4	85.6	800	9	CL457009	ZMMBB048	190	15	83.3	602	8	AZ95079
118	15.4	85.6	812	7	CK769679	957749.MA	191	15	83.3	603	2	BE039065
119	15.4	85.6	824	7	BZ995959	PUPJC70TD	192	15	83.3	609	5	BM473143
120	15.4	85.6	848	9	CC695247	OGVET81TV	193	15	83.3	609	8	BZ633986
121	15.4	85.6	859	8	BZ995953	PUPJC70TB	194	15	83.3	612	9	BX185625
122	15.4	85.6	871	9	CR266503	Forward.6	195	15	83.3	615	8	BM465633
123	15.4	85.6	912	9	CL298060	ZMMBB008	196	15	83.3	618	8	BZ679894
124	15.4	85.6	917	9	CL445921	ZMMBB046	197	15	83.3	620	6	CA190996
125	15.4	85.6	925	9	CL464318	SA1L_1219	198	15	83.3	622	2	BE214881
126	15.4	85.6	949	2	BF308719	601890011	199	15	83.3	632	5	BQ859935
127	15.4	85.6	994	2	CG326682	OGVCM87TV	200	15	83.3	632	1	BU992061
128	15.4	85.6	1016	4	BM048212	603620189	201	15	83.3	643	5	AV992458
129	15.4	85.6	1033	9	CNS06Y50	end.of	202	15	83.3	650	6	CB883667
130	15.4	85.6	1062	9	CNS02UTS	Al214921	203	15	83.3	654	5	BO859638
131	15.4	85.6	1141	7	CN970619	18116_123	204	15	83.3	654	6	CA090403
132	15.4	85.6	1292	2	BF967946	602827830	205	15	83.3	674	5	BM421513
133	15.4	85.6	1312	9	CG749006	P043-1-DO	206	15	83.3	681	5	BM419828
134	15.4	85.6	1461	2	BR971547	602239772	207	15	83.3	682	8	BM425288
135	15.2	84.4	1952	5	BK400921	BM400921	208	15	83.3	685	8	BZ636291
136	15	83.3	149	1	AV355231	AV355231	209	15	83.3	685	5	CL167604
137	15	83.3	273	5	BP636527	BP636527	210	15	83.3	688	5	BM020278
138	15	83.3	289	5	AV880807	AV880807	211	15	83.3	692	7	CO490450
139	15	83.3	337	6	CD479381	eca01-21m	212	15	83.3	696	8	BM935232
140	15	83.3	351	5	BY302652	BY302652	213	15	83.3	698	8	BM934884
141	15	83.3	360	5	BP629145	BP629145	214	15	83.3	706	6	CD867400
142	15	83.3	367	5	BM072723	BM072723	215	15	83.3	708	6	CA140075
143	15	83.3	367	5	BM074459	BM074459	216	15	83.3	717	8	BZ408702
144	15	83.3	383	9	CC871300	NDL_121P1	217	15	83.3	721	6	CE235788
145	15	83.3	388	9	CL281731	ZMMBB062	218	15	83.3	721	7	CR285094
146	15	83.3	390	9	CG904708	ZMMBB051	219	15	83.3	723	9	CC874884
147	15	83.3	392	8	B35497	HS-1029-A1-	220	15	83.3	727	1	AJ803466
148	15	83.3	393	5	BP639218	BP639218	221	15	83.3	730	8	BZ979822
149	15	83.3	396	1	AV784347	AV784347	222	15	83.3	736	5	BU001781
150	15	83.3	407	5	BP587928	BP587928	223	15	83.3	743	9	CL745885
151	15	83.3	413	5	BP633053	BP633053	224	15	83.3	747	6	CA266752
152	15	83.3	423	5	BM208153	BM208153	225	15	83.3	747	8	BZ467270
153	15	83.3	423	5	BM209971	BM209971	226	15	83.3	751	7	CO079632
154	15	83.3	425	1	AV819763	AV819763	227	15	83.3	753	5	BM117438
155	15	83.3	429	5	BP634162	BP634162	228	15	83.3	759	7	CK120352
156	15	83.3	439	8	BH719447	BOKKK04TR	229	15	83.3	760	8	BK455302
157	15	83.3	439	8	BH823447	BACPR17-H	230	15	83.3	761	7	CN180704
158	15	83.3	445	5	BP635711	BP635711	231	15	83.3	761	9	CL167603
159	15	83.3	448	5	BP593508	BP593508	232	15	83.3	762	9	AG468993
160	15	83.3	451	5	BP586034	BP586034	233	15	83.3	763	6	CA166813
161	15	83.3	474	1	AV879730	AV879730	234	15	83.3	764	9	BX163759
162	15	83.3	477	5	BY256777	BY256777	235	15	83.3	765	6	CA074505
163	15	83.3	480	5	AQ721815	HS_5560.B	236	15	83.3	769	5	BM407691
164	15	83.3	488	2	BP657167	OY2_21_E0	237	15	83.3	772	9	CC679148
165	15	83.3	489	2	BM004187	BM004187	238	15	83.3	778	8	CC093829
166	15	83.3	492	6	CB260357	51-E9624-	239	15	83.3	784	9	CC450880
167	15	83.3	494	1	AA854676	AJ63106.s	240	15	83.3	790	5	BM113557
168	15	83.3	501	7	CN617766	TGBSTYm3	241	15	83.3	793	8	BZ485455
169	15	83.3	504	2	AW744896	1G1_384.H	242	15	83.3	794	8	BZ406661
170	15	83.3	506	9	CR400606	ArabiIdops	243	15	83.3	803	9	CG230159

244	15	83.3	810	8	BH426750	BH426750	BOGHJ20TR	C 317	14.8	82.2	311	6	CF050330	OCM10h03.
245	15	83.3	816	9	CG265163	CG265163	OG0FAV1TH	C 318	14.8	82.2	311	6	CF055390	OCN5a03.Y
246	15	83.3	822	9	CC117600	CC117600	NDL.3A7.T	C 319	14.8	82.2	312	6	CF085819	QAN15a07.
247	15	83.3	834	8	BZ462557	BZ462557	BONEA51TF	C 320	14.8	82.2	312	6	CD986478	QAN22c11.
248	15	83.3	836	8	BZ746910	BZ746910	PIECET35TD	C 321	14.8	82.2	312	6	CD986562	QAN23d09.
249	15	83.3	838	8	CC335362	CC335362	OCUAF74TV	C 322	14.8	82.2	312	6	CD986589	QAN23g04.
250	15	83.3	839	8	BZ466556	BZ466556	BONNV86TR	C 323	14.8	82.2	312	6	CD987138	QAN7e11.Y
251	15	83.3	851	8	BH706797	BH706797	BOMDV28TF	C 324	14.8	82.2	312	6	CD989529	CAT3a12.Y
252	15	83.3	872	3	CNS0ADG2	CNS0ADG2	ArabiIdops	C 325	14.8	82.2	313	6	CF033172	QCF13d02.
253	15	83.3	877	9	CC658870	CC658870	OG0AN90TV	C 326	14.8	82.2	313	6	CA299790	SCMCLV103
254	15	83.3	893	9	CC682520	CC682520	OGUHO03TH	C 327	14.8	82.2	314	6	CD165028	ML1-.0093T
255	15	83.3	897	9	CC313186	CC313186	OGXAR33TV	C 328	14.8	82.2	315	6	CD997040	OBd2h02.x
256	15	83.3	902	9	CNS0170R	CNS0170R	Tetraodon	C 329	14.8	82.2	321	6	CD724442	MK.5.13.P
257	15	83.3	902	9	CG265171	CG265171	OG0FAV1TV	C 330	14.8	82.2	323	5	BH552601	952076C05
258	15	83.3	906	8	BZ786487	BZ786487	PUGDL02TD	C 331	14.8	82.2	325	8	BZ761265	4L4.B
259	15	83.3	907	9	CC658861	CC658861	OG0AN90TH	C 332	14.8	82.2	330	6	CD987818	QA05d08.Y
260	15	83.3	916	8	CC343433	CC343433	OGPBR37TV	C 333	14.8	82.2	330	6	CD987839	QA05f08.Y
261	15	83.3	928	9	CC6844197	CC6844197	NDL.132K8	C 334	14.8	82.2	330	6	CF046616	QCK27h12.Y
262	15	83.3	938	9	CNS01TVG	CNS01TVG	Tetraodon	C 335	14.8	82.2	330	6	CF049903	QCL4h12.Y
263	15	83.3	999	6	CA213451	CA213451	SCGGSB114	C 336	14.8	82.2	330	6	CF051213	OCM22f10.
264	15	83.3	1020	5	CNS072YK	CNS072YK	clone.BA0	C 337	14.8	82.2	330	6	CF061267	OCCT19f06.
265	15	83.3	1022	5	BH398891	BH398891	603534208	C 338	14.8	82.2	330	6	CF061497	OCCT12e02.
266	15	83.3	1101	9	CNS00009	CNS00009	Drosophi1	C 339	14.8	82.2	330	6	CF061902	QCT2f06.Y
267	15	83.3	1113	9	CNS0600Z	CNS0600Z	T3 end of	C 340	14.8	82.2	330	6	CF062920	OCU5c08.Y
268	15	83.3	1515	3	CF756560	CF756560	DSAF1.7.A	C 341	14.8	82.2	331	6	CF063732	ETH1.12.B
269	14.8	82.2	141	7	BZ290873	BZ290873	ArabiIdops	C 342	14.8	82.2	331	6	CD431949	OCU7a10.Y
270	14.8	82.2	147	7	AX290873	AX290873	ArabiIdops	C 343	14.8	82.2	332	6	CF039400	OCH36a12.
271	14.8	82.2	155	7	AI901401	AI901401	618003D04	C 344	14.8	82.2	332	6	CF057339	OC031d05.
272	14.8	82.2	156	7	CO296379	CO296379	EX213520.	C 345	14.8	82.2	332	6	CF061957	OCCT3c11.Y
273	14.8	82.2	173	2	AN017653	AN017653	614061D08	C 346	14.8	82.2	332	6	CF062758	OCU3c09.Y
274	14.8	82.2	192	9	CG319258	CG319258	OG1CT74TH	C 347	14.8	82.2	332	6	CF064138	OCU7f07.Y
275	14.8	82.2	212	1	AV931909	AV931909	Av931909	C 348	14.8	82.2	332	6	CF064176	OCUB06.Y
276	14.8	82.2	216	6	AB082335	AB082335	Drosophi1	C 349	14.8	82.2	333	7	CO096789	GR_Ea20E
277	14.8	82.2	245	6	CB489366	CB489366	omyKrcDh0	C 350	14.8	82.2	333	9	TA89B06P	T. Ducei
278	14.8	82.2	249	9	AN0055390	AN0055390	614080C11	C 351	14.8	82.2	333	6	CA004073	HS16H9T
279	14.8	82.2	249	9	CC7966146	CC7966146	SALR.0964	C 352	14.8	82.2	336	6	CF050679	OCM15h08.
280	14.8	82.2	252	6	CF063520	CF063520	QC022e05.	C 353	14.8	82.2	340	6	CF050679	qea1e07.s
281	14.8	82.2	257	6	CD989883	CD989883	QAU4d01.Y	C 354	14.8	82.2	343	1	AI193389	CB231982
282	14.8	82.2	264	5	BU065667	BU065667	Fgr.8.C24	C 355	14.8	82.2	346	5	B0163294	95208bD10
283	14.8	82.2	267	6	CF115617	CF115617	SSES030N2	C 356	14.8	82.2	349	5	BO163294	95208bD10
284	14.8	82.2	276	2	BF488015	BF488015	AI23259.5	C 357	14.8	82.2	349	6	CA252519	MTU7TL.P1
285	14.8	82.2	279	2	BH303871	BH303871	BR038571	C 358	14.8	82.2	350	6	AI1795703	614001P06
286	14.8	82.2	286	6	CB081591	CB081591	hk53h06.G	C 359	14.8	82.2	350	6	CF061192	CL101f901.
287	14.8	82.2	286	6	BI059597	BI059597	MR4-UT005	C 360	14.8	82.2	350	6	CF028682	OCCT18S01.
288	14.8	82.2	290	4	BI059597	BI059597	MR4-UT005	C 361	14.8	82.2	352	6	CF028682	OCCT18S02.
289	14.8	82.2	292	6	CD436372	CD436372	E101N0352	C 362	14.8	82.2	354	5	BU062033	OCCT13h02.
290	14.8	82.2	293	9	CG685418	CG685418	ZMMBBO15	C 363	14.8	82.2	354	6	CD988118	BO062033
291	14.8	82.2	296	9	CG097332	CG097332	PUIFN61TB	C 364	14.8	82.2	355	2	AM126723	FCR.1.E04
292	14.8	82.2	297	6	CG962772	CG962772	SDR.132.G	C 365	14.8	82.2	356	4	AI172417	SA14f01.Y
293	14.8	82.2	298	5	BU571931	BU571931	946166D01	C 366	14.8	82.2	359	1	BG517362	SH1Z2.7.A
294	14.8	82.2	299	5	CB381663	CB381663	3529.1.10	C 367	14.8	82.2	360	7	D69723	CEK773C1F
295	14.8	82.2	300	6	CB488899	CB488899	omyKrcDh0	C 368	14.8	82.2	363	7	CO300168	947061E05
296	14.8	82.2	300	9	CG097336	CG097336	PUIFN61TD	C 369	14.8	82.2	364	5	BP068550	CEK177523.
297	14.8	82.2	301	1	AV546446	AV546446	AV546446	C 370	14.8	82.2	368	8	BZ777709	1142804.9
298	14.8	82.2	302	5	BQ163312	BQ163312	952080F11	C 371	14.8	82.2	369	6	CF057498	OC033g06.
299	14.8	82.2	302	6	BM985630	BM985630	952076C05	C 372	14.8	82.2	370	6	CF056439	OC015f08.
300	14.8	82.2	307	4	BM985630	BM985630	952076C05	C 373	14.8	82.2	372	7	CF400558	RC019f06.
301	14.8	82.2	310	8	CC422313	CC422313	PUIHEB22TD	C 374	14.8	82.2	375	6	CD954718	RTW11.T6.C
302	14.8	82.2	311	6	CD970550	CD970550	QAD19C06.	C 375	14.8	82.2	375	6	CD962221	SDN.141.G
303	14.8	82.2	311	6	CD985763	CD985763	QAN14d01.	C 376	14.8	82.2	375	6	CD964653	SED.6.Gen
304	14.8	82.2	311	6	CD986140	CD986140	QAN19B12.	C 377	14.8	82.2	376	4	BG241792	SR7.7.Ge
305	14.8	82.2	311	6	CD986276	CD986276	QAN19B12.Y	C 378	14.8	82.2	376	4	AG261131	PH1Z2.50.
306	14.8	82.2	311	6	CD986416	CD986416	QAN21e09.	C 379	14.8	82.2	376	6	CD945524	RDZ.71.Ge
307	14.8	82.2	311	6	CD987053	CD987053	QAN6d12.Y	C 380	14.8	82.2	377	9	AG261131	Letus.cor
308	14.8	82.2	311	6	CD987060	CD987060	QAN6e09.Y	C 381	14.8	82.2	378	5	BE554034	BP654034
309	14.8	82.2	311	6	CD987465	CD987465	QAO1b05.Y	C 382	14.8	82.2	379	2	BE552696	946083H11
310	14.8	82.2	311	6	CD987747	CD987747	QAO4e12.Y	C 383	14.8	82.2	380	5	CF117128	fp264.z1
311	14.8	82.2	311	6	CD988423	CD988423	QAP4a01.Y	C 384	14.8	82.2	381	5	BU051404	111043A0
312	14.8	82.2	311	6	CD988556	CD988556	QAO1906.Y	C 385	14.8	82.2	382	5	BQ133666	952080F11
313	14.8	82.2	311	6	CD989599	CD989599	QAT4a11.Y	C 386	14.8	82.2	385	5	BO133666	952080F11
314	14.8	82.2	311	6	CD989622	CD989622	QAT4d10.Y	C 387	14.8	82.2	387	6	CF040826	OCN89787
315	14.8	82.2	311	6	CD989655	CD989655	QAV4d12.Y	C 388	14.8	82.2	389	6	CF036817	OC119a03.
316	14.8	82.2	311	6	CF035317	CF035317	QCG36h01.	C 389	14.8	82.2	389	6	CF057774	OC036g09.

C 390	14.8	82.2	390	6	CF057476	QCO32603.	463	14.8	82.2	491	5	CA895507	CA895507
C 391	14.8	82.2	394	8	B2420459	1E6804.b	464	14.8	82.2	492	4	BA103396	BG103396
C 392	14.8	82.2	397	7	CF624454	zmrw05.0	465	14.8	82.2	495	4	BG509546	BG509546
C 393	14.8	82.2	399	1	A1944057	614034D09	466	14.8	82.2	496	6	CF023482	CF023482
C 394	14.8	82.2	403	6	CB082250	hK65909.5	467	14.8	82.2	497	7	CK907743	CK907743
C 395	14.8	82.2	408	7	CF757962	DSAP1.25	468	14.8	82.2	498	7	AC253113	AC253113
C 396	14.8	82.2	411	5	BO410784	GA_EB003	469	14.8	82.2	499	7	CN131785	CN131785
C 397	14.8	82.2	412	5	BP669266	BP669266	470	14.8	82.2	499	4	BT610381	BT610381
C 398	14.8	82.2	414	1	AV803597	AV803597	471	14.8	82.2	499	6	CF043032	CF043032
C 399	14.8	82.2	415	1	A1967286	A1967286	472	14.8	82.2	501	1	A1948084	A1948084
C 400	14.8	82.2	415	1	AV807784	AV807784	473	14.8	82.2	501	6	CA898057	CA898057
C 401	14.8	82.2	415	5	BO410783	GA_EB003	474	14.8	82.2	501	6	CF020279	CF020279
C 402	14.8	82.2	417	6	CA225572	SCRLFL06	475	14.8	82.2	502	1	AV780692	AV780692
C 403	14.8	82.2	418	8	BP669638	SP_BA007	476	14.8	82.2	504	1	BM325647	BM325647
C 404	14.8	82.2	420	4	BP459581	BT659381	477	14.8	82.2	505	4	BM338464	BM338464
C 405	14.8	82.2	425	2	BT251189	EST418449	478	14.8	82.2	505	4	BO663743	BO663743
C 406	14.8	82.2	425	6	CB490711	OMYKTB10	479	14.8	82.2	505	6	CB334351	CB334351
C 407	14.8	82.2	430	7	CN487063	EST1709.P	480	14.8	82.2	505	6	CB380651	CB380651
C 408	14.8	82.2	431	1	AV540666	AV540666	481	14.8	82.2	505	7	CO325585	CO325585
C 409	14.8	82.2	431	1	AV927254	AV927254	482	14.8	82.2	506	1	A1901384	A1901384
C 410	14.8	82.2	431	6	CD996528	QBC4h02.x	483	14.8	82.2	506	1	BM551605	BM551605
C 411	14.8	82.2	432	2	BR470740	UT-M-BH3-	484	14.8	82.2	506	8	A2769746	A2769746
C 412	14.8	82.2	434	2	AM017490	614059F06	485	14.8	82.2	506	9	CL719019	CL719019
C 413	14.8	82.2	435	2	AM065436	614051G11	486	14.8	82.2	508	2	BF502367	BF502367
C 414	14.8	82.2	435	2	AM179467	618046B03	487	14.8	82.2	509	6	CD725068	CD725068
C 415	14.8	82.2	436	5	BP049179	BP049179	488	14.8	82.2	510	5	BU694536	BU694536
C 416	14.8	82.2	440	5	BE598324	PI1_B1_H0	489	14.8	82.2	510	7	CO139242	CO139242
C 417	14.8	82.2	440	5	BY265902	BY265902	490	14.8	82.2	511	7	CF796943	CF796943
C 418	14.8	82.2	442	7	CO300457	EX778231.	491	14.8	82.2	512	5	CF019849	CF019849
C 419	14.8	82.2	444	4	BM061072	KS01026B0	492	14.8	82.2	512	6	CF019849	CF019849
C 420	14.8	82.2	444	6	CA011441	HT051070	493	14.8	82.2	513	4	BG509559	BG509559
C 421	14.8	82.2	445	4	BA484585	BJ484585	494	14.8	82.2	513	4	BI192013	BI192013
C 422	14.8	82.2	447	2	BF500733	AT15745.5	495	14.8	82.2	513	5	BP040284	BP040284
C 423	14.8	82.2	449	2	BE225135	946023G07	496	14.8	82.2	514	7	CO192461	CO192461
C 424	14.8	82.2	449	4	BM073165	MEST61-G0	497	14.8	82.2	514	4	BM340599	BM340599
C 425	14.8	82.2	450	5	BO658867	HD03C020	498	14.8	82.2	516	1	A2434965	A2434965
C 426	14.8	82.2	450	5	BU063286	FY2_016	499	14.8	82.2	516	2	BP277188	BP277188
C 427	14.8	82.2	455	4	BU451878	BU451878	500	14.8	82.2	517	2	AM331782	AM331782
C 428	14.8	82.2	455	8	CC085614	CSU-K33r.	501	14.8	82.2	517	7	CE634551	CE634551
C 429	14.8	82.2	456	2	AM416853	52586.MAR	502	14.8	82.2	518	7	CO144784	CO144784
C 430	14.8	82.2	456	7	CN107405	EC2CA30A	503	14.8	82.2	519	1	A1629677	A1629677
C 431	14.8	82.2	457	4	BI096490	949017H03	504	14.8	82.2	520	4	BM072942	BM072942
C 432	14.8	82.2	457	7	CO183533	EC05920.5	505	14.8	82.2	521	4	BM072939	BM072939
C 433	14.8	82.2	460	5	BU572174	946168F01	506	14.8	82.2	521	8	BZ484234	BZ484234
C 434	14.8	82.2	460	5	CF007488	QB12908.P	507	14.8	82.2	521	9	CG893876	CG893876
C 435	14.8	82.2	461	7	CF007488	QB12908.X	508	14.8	82.2	525	5	BO534739	BO534739
C 436	14.8	82.2	461	7	CO284869	EX167134.	509	14.8	82.2	525	9	CG211198	CG211198
C 437	14.8	82.2	462	2	AM282936	LG1_304.D	510	14.8	82.2	526	5	BO655331	BO655331
C 438	14.8	82.2	466	6	CA891815	BO168B07-	511	14.8	82.2	526	6	CD724989	CD724989
C 439	14.8	82.2	467	4	BT174529	OSFR047H7	512	14.8	82.2	527	4	BI097741	BI097741
C 440	14.8	82.2	468	5	BE2964342	1091028D0	513	14.8	82.2	528	4	BO642533	BO642533
C 441	14.8	82.2	469	6	CD724930	MX_16_24	514	14.8	82.2	528	6	CA005388	CA005388
C 442	14.8	82.2	472	8	BR854138	SALX_0787	515	14.8	82.2	529	1	A1833441	A1833441
C 443	14.8	82.2	473	6	CD997375	QBD6F05.X	516	14.8	82.2	529	6	CA890751	CA890751
C 444	14.8	82.2	475	1	AV923725	AV923725	517	14.8	82.2	529	6	CB179706	CB179706
C 445	14.8	82.2	476	6	CB815392	3529.1.73	518	14.8	82.2	530	5	BO143389	BO143389
C 446	14.8	82.2	479	1	A1947693	603026G08	519	14.8	82.2	530	5	BU695934	BU695934
C 447	14.8	82.2	480	5	BU811359	UL837C10	520	14.8	82.2	530	6	CA1144010	CA1144010
C 448	14.8	82.2	481	2	BE633388	946024E09	521	14.8	82.2	530	6	CB395975	CB395975
C 449	14.8	82.2	481	7	CO140775	EST833446	522	14.8	82.2	531	2	BF487001	BF487001
C 450	14.8	82.2	484	1	A1455485	L24478.3	523	14.8	82.2	531	6	CF042088	CF042088
C 451	14.8	82.2	484	5	BO539325	MEST614-D	524	14.8	82.2	532	4	AO666342	AO666342
C 452	14.8	82.2	484	7	CF636693	zmrw00.0	525	14.8	82.2	533	4	BO360645	BO360645
C 453	14.8	82.2	484	7	CO136957	EST833628	526	14.8	82.2	533	6	CF005217	CF005217
C 454	14.8	82.2	484	7	AO602487	HS_218.A	527	14.8	82.2	533	6	CB833657	CB833657
C 455	14.8	82.2	485	6	CA06213	HU05D02u	528	14.8	82.2	535	2	AM090922	AM090922
C 456	14.8	82.2	487	2	AM061663	660012G11	529	14.8	82.2	535	2	AM267411	AM267411
C 457	14.8	82.2	488	2	BE510791	946055E05	530	14.8	82.2	537	5	BO103409	BO103409
C 458	14.8	82.2	488	2	BE513397	946070E05	531	14.8	82.2	537	6	CA485001	CA485001
C 459	14.8	82.2	489	7	CF796142	892370.MA	532	14.8	82.2	537	7	CK441039	CK441039
C 460	14.8	82.2	490	7	BU063191	FYR_2.Y16	533	14.8	82.2	538	6	CB605594	CB605594
C 461	14.8	82.2	491	5	BO618922	RNO5E01D0	534	14.8	82.2	539	1	A0243024	A0243024
C 462	14.8	82.2	491	5	BO618922	RNO5E01D0	535	14.8	82.2	539	1	A0243024	A0243024

C 536	14.8	82.2	539	4	BM334338	MEST135-G	609	14.8	82.2	587	6	CB815393	3529_1_73
C 537	14.8	82.2	539	6	CA000650	HS07F02u	610	14.8	82.2	587	7	CO143124	EST857795
C 538	14.8	82.2	539	6	CF004068	QBH26C03.	611	14.8	82.2	589	1	AI516117	LD4873.5
C 539	14.8	82.2	540	2	AM067459	614044B08	612	14.8	82.2	589	6	CD331173	St-Pu537.
C 540	14.8	82.2	541	6	CD987103	CD987103	613	14.8	82.2	590	4	BU176410	BU176410
C 541	14.8	82.2	543	2	BE975896	bsa5h01.x	614	14.8	82.2	590	6	CA365765	CA365765
C 542	14.8	82.2	544	5	BM697931	LI21h1102	615	14.8	82.2	590	8	AO609236	AO609236
C 543	14.8	82.2	544	6	CA277725	CA277725	616	14.8	82.2	591	4	BM641879	BM641879
C 544	14.8	82.2	544	6	CA277725	CA277725	617	14.8	82.2	591	6	CF021127	CF021127
C 545	14.8	82.2	545	6	CA005899	CA005899	618	14.8	82.2	592	4	BM355249	BM355249
C 546	14.8	82.2	545	6	BC034095	1091002A0	619	14.8	82.2	593	4	BM337500	BM337500
C 547	14.8	82.2	549	8	BZ704236	PUBMD51TD	620	14.8	82.2	594	1	AI258424	AI258424
C 548	14.8	82.2	550	2	BE509832	946051F10	621	14.8	82.2	594	4	BG458580	BG458580
C 549	14.8	82.2	550	4	BT674682	949066B02	622	14.8	82.2	596	4	BM067312	BM067312
C 550	14.8	82.2	550	6	CA297567	SCCSPD02	623	14.8	82.2	596	7	CK368676	CK368676
C 551	14.8	82.2	551	4	BM465496	RH1722_45	624	14.8	82.2	597	1	AV924580	AV924580
C 552	14.8	82.2	551	4	BM340362	MEST328-D	625	14.8	82.2	597	6	CF075465	EST1282 Z
C 553	14.8	82.2	551	7	CK827923	zmrw05_0	626	14.8	82.2	598	2	BE975897	BE975897
C 554	14.8	82.2	552	7	COS28581	3530_1_18	627	14.8	82.2	598	6	CA831697	CA831697
C 555	14.8	82.2	553	2	BE510298	946051F10	628	14.8	82.2	599	5	BU698481	BU698481
C 556	14.8	82.2	554	4	BM482692	BT482692	629	14.8	82.2	599	6	CB638851	CB638851
C 557	14.8	82.2	554	4	CB488935	OMYKtchDh	630	14.8	82.2	600	6	CB834054	CB834054
C 558	14.8	82.2	554	7	CO224020	WS01020.B	631	14.8	82.2	600	6	CD725045	CD725045
C 559	14.8	82.2	556	4	BI894058	sa159e08.	632	14.8	82.2	601	1	AI063268	AI063268
C 560	14.8	82.2	557	1	AI998458	701545756	633	14.8	82.2	601	6	CA111154	CA111154
C 561	14.8	82.2	557	4	BI992601	1020064H0	634	14.8	82.2	601	6	CA229375	CA229375
C 562	14.8	82.2	557	4	BM333149	MEST184-F	635	14.8	82.2	601	7	CO049108	CO049108
C 563	14.8	82.2	557	4	BM870756	mrm011x0	636	14.8	82.2	603	7	CO480761	CO480761
C 564	14.8	82.2	558	2	AM682742	EST01314	637	14.8	82.2	604	5	CD998008	CD998008
C 565	14.8	82.2	559	5	BU079733	946147A08	638	14.8	82.2	605	2	BF255168	BF255168
C 566	14.8	82.2	561	4	BG458421	947045H08	639	14.8	82.2	605	2	BF424109	BF424109
C 567	14.8	82.2	562	6	CA627468	w1n.Pk15	640	14.8	82.2	605	5	BU060744	BU060744
C 568	14.8	82.2	562	6	CB885594	3529_1_88	641	14.8	82.2	605	5	BU571932	BU571932
C 569	14.8	82.2	563	5	BU048932	109100Q0H0	642	14.8	82.2	605	6	CF013793	CF013793
C 570	14.8	82.2	563	7	CK972419	4102879.B	643	14.8	82.2	606	6	CA122263	CA122263
C 571	14.8	82.2	565	4	BM348446	MEST291-A	644	14.8	82.2	606	6	CB885593	CB885593
C 572	14.8	82.2	566	1	AI292503	GH15425.5	645	14.8	82.2	606	6	CF017522	CF017522
C 573	14.8	82.2	566	6	CA127331	SCCCLRC0	646	14.8	82.2	607	8	AO249624	AO249624
C 574	14.8	82.2	568	6	CA001227	HS18N03u	647	14.8	82.2	608	6	CA155905	CA155905
C 575	14.8	82.2	568	6	CB329843	3529_1_24	648	14.8	82.2	609	4	BG048071	BG048071
C 576	14.8	82.2	568	6	CB833536	3529_1_81	649	14.8	82.2	609	6	CB885428	CB885428
C 577	14.8	82.2	569	8	AZ842963	2M0203G22	650	14.8	82.2	610	1	AI295026	AI295026
C 578	14.8	82.2	570	5	BQ115117	EX176419.	651	14.8	82.2	610	4	BM388031	BM388031
C 579	14.8	82.2	572	4	BM337899	MEST218-H	652	14.8	82.2	610	4	BM079123	BM079123
C 580	14.8	82.2	573	2	BE948881	UI-M-BH3	653	14.8	82.2	610	6	CB351716	CB351716
C 581	14.8	82.2	573	4	BM332193	MEST153-F	654	14.8	82.2	610	6	CF014172	CF014172
C 582	14.8	82.2	573	6	CB251926	3529_1_19	655	14.8	82.2	611	4	BG048223	BG048223
C 583	14.8	82.2	574	4	BT498933	sgg17611.	656	14.8	82.2	611	7	CN148227	CN148227
C 584	14.8	82.2	574	6	CB411334	3529_1_60	657	14.8	82.2	612	4	BM383577	BM383577
C 585	14.8	82.2	575	2	BE364692	PT1_14.C0	658	14.8	82.2	612	5	BM882147	BM882147
C 586	14.8	82.2	575	4	BM350429	MEST265-H	659	14.8	82.2	612	6	CB380753	CB380753
C 587	14.8	82.2	575	6	CD204185	HS1_4.G08	660	14.8	82.2	612	6	CB491555	CB491555
C 588	14.8	82.2	577	8	AO303486	HS_3237_B	661	14.8	82.2	612	6	CB833656	CB833656
C 589	14.8	82.2	578	6	CF006455	OBT1905-X	662	14.8	82.2	612	8	AO613796	AO613796
C 590	14.8	82.2	579	6	CA222951	SCBGRFL05	663	14.8	82.2	613	4	BI227970	BI227970
C 591	14.8	82.2	580	1	AI649965	AL649965	664	14.8	82.2	614	2	BF485187	BF485187
C 592	14.8	82.2	580	1	BQ107109	NXLV087_A	665	14.8	82.2	615	4	BM341708	BM341708
C 593	14.8	82.2	581	2	BE344896	946029G08	666	14.8	82.2	615	6	CA989180	CA989180
C 594	14.8	82.2	581	6	CF020328	QBN4d12.x	667	14.8	82.2	616	4	BM347767	BM347767
C 595	14.8	82.2	582	4	BT232564	RE28649-H	668	14.8	82.2	616	4	BM348754	BM348754
C 596	14.8	82.2	582	4	BM339108	MEST236-H	669	14.8	82.2	618	7	CK605419	CK605419
C 597	14.8	82.2	582	4	CD990509	QAZ2A08.Y	670	14.8	82.2	619	4	BT551129	BT551129
C 598	14.8	82.2	582	6	CF613654	CS5007943	671	14.8	82.2	619	6	CF002276	CF002276
C 599	14.8	82.2	583	7	BE845486	233293.BA	672	14.8	82.2	619	2	AM671487	AM671487
C 600	14.8	82.2	583	4	BI636542	SD18251.5	673	14.8	82.2	620	7	BI486215	BI486215
C 601	14.8	82.2	583	4	BO662418	HS02E20u	674	14.8	82.2	620	6	CD527756	CD527756
C 602	14.8	82.2	583	6	CA229329	SCAGFL302	675	14.8	82.2	620	9	CL824677	CL824677
C 603	14.8	82.2	583	6	CA829341	3529_1_3	676	14.8	82.2	621	2	AM498332	AM498332
C 604	14.8	82.2	584	6	CA108208	SCGBHRT05	677	14.8	82.2	621	5	BU696640	BU696640
C 605	14.8	82.2	584	6	CO147023	ESTB22076	678	14.8	82.2	621	5	BU698735	BU698735
C 606	14.8	82.2	585	4	BM350393	MEST265-C	679	14.8	82.2	621	6	AZ009763	AZ009763
C 607	14.8	82.2	585	4	CB240059	3529_1_15	680	14.8	82.2	621	6	CA113858	CA113858
C 608	14.8	82.2	586	6			681	14.8	82.2	622	6		

682	14.8	82.2	622	6	CF004596	QBH5A07.X	755	14.8	82.2	666	5	BU065758	Fgr_8_G22
683	14.8	82.2	623	6	B5599343	P11_88.D0	756	14.8	82.2	667	7	CK755616	par01-2ms
684	14.8	82.2	623	6	CA282084	SCAGSD204	757	14.8	82.2	667	7	CF624183	zmrw05.0
685	14.8	82.2	623	6	CA988931	3529.1.11	758	14.8	82.2	667	7	CF630966	zmrw04.0
686	14.8	82.2	623	7	CO049056	db05f05.	759	14.8	82.2	668	4	BI487191	RE70710.5
687	14.8	82.2	623	8	CH669551	BOMLE94TF	760	14.8	82.2	668	5	BU065958	Fgr_6_P13
688	14.8	82.2	623	9	CL700581	SP_BA006	761	14.8	82.2	668	7	CK134651	RE70710.3
689	14.8	82.2	625	4	BM332708	MEST176-E	762	14.8	82.2	669	4	BI959787	HVSMET002
690	14.8	82.2	625	5	BU079705	946146G08	763	14.8	82.2	669	9	CG377824	CG377824
691	14.8	82.2	625	6	CF007122	QBK4C04.X	764	14.8	82.2	670	2	BF251484	EST148745
692	14.8	82.2	626	6	CF001395	QBK4C04.X	765	14.8	82.2	670	4	BC842553	MEST33-C0
693	14.8	82.2	626	7	CK133793	RE38475.3	766	14.8	82.2	670	9	CR487848	Medicago
694	14.8	82.2	627	1	AL731320	BNLGH1921	767	14.8	82.2	671	5	BU060768	Fgr-C_1.J
695	14.8	82.2	627	6	CF004942	QBH7C11.X	768	14.8	82.2	671	9	CG118904	PULSG6TB
696	14.8	82.2	628	6	CF003636	QBH21F06.	769	14.8	82.2	672	4	BC842564	MEST33-D0
697	14.8	82.2	629	1	AI388100	GHI1883.5	770	14.8	82.2	672	7	CF632846	zmrw04.0
698	14.8	82.2	629	6	CA112029	SCCCLB102	771	14.8	82.2	673	5	BU062831	Fgr_2_E24
699	14.8	82.2	629	6	CA486423	WHE3435.E	772	14.8	82.2	674	2	BF251401	EST148662
700	14.8	82.2	629	6	CE079821	hp77C10.D	773	14.8	82.2	674	7	CK135442	OX1_32.G0
701	14.8	82.2	632	6	CE688195	CHEST-10-F	774	14.8	82.2	676	4	BI577718	RE7144T.5
702	14.8	82.2	632	7	CK661212	LP20456.5	775	14.8	82.2	676	4	BM334216	MEST134-B
703	14.8	82.2	633	5	BO291309	NKRVO58.E	776	14.8	82.2	677	5	BU699234	LL21n1269
704	14.8	82.2	634	6	CD460615	F909.01d1	777	14.8	82.2	678	1	AI650105	AEWYAM85
705	14.8	82.2	635	2	BB625822	BB625822	778	14.8	82.2	678	6	CD462423	ETH1_37.F
706	14.8	82.2	635	5	BU042766	PR_LBA001	779	14.8	82.2	678	7	CK978064	4109386.B
707	14.8	82.2	635	6	CE171494	PK0602601	780	14.8	82.2	679	6	CA806516	ESG013d.E
708	14.8	82.2	636	2	AW288963	660015G02	781	14.8	82.2	679	9	CE889028	tigr-gs8-
709	14.8	82.2	637	1	AI657528	AEMTBD32	782	14.8	82.2	681	7	CO521424	3530_1.14
710	14.8	82.2	637	4	BM329776	PIC1_43.F	783	14.8	82.2	681	7	CO530705	3530_1.20
711	14.8	82.2	638	9	CE630673	tigr-gs8-	784	14.8	82.2	683	5	BM883932	BM883932
712	14.8	82.2	639	7	BE394545	601310555	785	14.8	82.2	683	6	CA236267	SCMCFLE501
713	14.8	82.2	639	7	CK369883	zmrw0485	786	14.8	82.2	683	7	CK973943	4104695.B
714	14.8	82.2	640	6	CA229997	SCUPFL3C0	787	14.8	82.2	683	9	AG182221	Pan_trog1
715	14.8	82.2	641	1	AI728488	BNLGH1108	788	14.8	82.2	685	5	BY704685	BY704685
716	14.8	82.2	641	2	BF488088	AT23172.5	789	14.8	82.2	685	6	CA261088	SCCCLB200
717	14.8	82.2	641	5	CD526392	EST1821.Z	790	14.8	82.2	686	1	AIU18613	AIU18613
718	14.8	82.2	642	6	BO410976	GA_ED003	791	14.8	82.2	686	4	BM358135	BM358135
719	14.8	82.2	642	9	CL703061	SP_BA009	792	14.8	82.2	687	4	BM338300	MESTY224-SQ
720	14.8	82.2	643	8	AC953905	2A953905	793	14.8	82.2	687	5	BO766149	EBX008-SQ
721	14.8	82.2	643	8	BE642544	MEST33-B0	794	14.8	82.2	688	4	BM698080	MEST375-D
722	14.8	82.2	644	4	BI214977	RE21150.5	795	14.8	82.2	688	4	BM99967	ca18d04.
723	14.8	82.2	645	4	BM337075	MEST202-H	796	14.8	82.2	689	2	BF276172	GA_ED002
724	14.8	82.2	645	6	CA298342	SCRFL803	797	14.8	82.2	690	4	BM338765	BM338765
725	14.8	82.2	645	7	CA148301	WOUNDI_55	798	14.8	82.2	690	4	BM351687	MEST321-H
726	14.8	82.2	646	5	EX878932	EX878932	799	14.8	82.2	691	4	BC637486	Zml0_10e0
727	14.8	82.2	647	4	BI579546	RE73592.5	800	14.8	82.2	691	6	CA288330	SCUFL301
728	14.8	82.2	647	5	BU060077	Fgr-C_0.K	801	14.8	82.2	692	4	BM306574	BM306574
729	14.8	82.2	649	9	CB613747	c1gr-gs8-	802	14.8	82.2	692	5	CA290134	CA290134
730	14.8	82.2	649	9	AM120344	614089C12	803	14.8	82.2	692	6	BJ111889	BJ111889
731	14.8	82.2	650	2	CL785314	OR_BBA010	804	14.8	82.2	693	4	CA289543	SCAGFL800
732	14.8	82.2	650	9	CA148301	MEST395-F	805	14.8	82.2	694	6	CA202228	SCSBFL103
733	14.8	82.2	651	4	BU064549	Fgr_5_C14	806	14.8	82.2	694	7	CF669689	RTCNT1_45
734	14.8	82.2	652	5	BU064549	Fgr_5_C14	807	14.8	82.2	694	7	BC936617	Zm08_04b0
735	14.8	82.2	653	7	CA135371	OR_3F_G0	808	14.8	82.2	695	4	CA094351	SCCCLB400
736	14.8	82.2	654	8	BE677854	PUBBGS7TD	809	14.8	82.2	695	6	AG182317	Pan_trog1
737	14.8	82.2	655	7	CO530706	3530_1.20	810	14.8	82.2	695	9	CA446346	WOUNDI_56
738	14.8	82.2	656	5	BU062210	Fgr_1_H21	811	14.8	82.2	696	6	CN148382	CN148382
739	14.8	82.2	656	7	CN881248	010425AAS	812	14.8	82.2	696	7	BM774387	CH230_35D
740	14.8	82.2	657	6	CD208455	H51_37_C1	813	14.8	82.2	697	8	BM774387	CH230_35D
741	14.8	82.2	658	4	BE841971	MEST33-C0	814	14.8	82.2	697	8	BM774387	CH230_35D
742	14.8	82.2	658	4	BE842131	MEST36-E1	815	14.8	82.2	697	7	BU065866	Fgr_8_L12
743	14.8	82.2	659	4	BM074759	MEST297-D0	816	14.8	82.2	698	5	BU065866	Fgr_8_L12
744	14.8	82.2	660	4	BC874190	MEST47-DO	817	14.8	82.2	698	7	CK761162	acti02-Sms
745	14.8	82.2	660	6	CB486596	omykrbho	818	14.8	82.2	699	6	CF273034	EST19596.Z
746	14.8	82.2	660	6	CL827350	OR_CBA004	819	14.8	82.2	700	6	CF006279	QBI19D08.
747	14.8	82.2	661	5	BU064367	Fgr_4_K22	820	14.8	82.2	701	4	BM629384	BM629384
748	14.8	82.2	661	6	CA243864	SCCFL407	821	14.8	82.2	701	4	BM439781	MESTY256-B
749	14.8	82.2	663	4	BM348162	MEST287-C	822	14.8	82.2	702	6	CA111363	SCAGLB106
750	14.8	82.2	663	5	BE623048	USDA-FP-0	823	14.8	82.2	704	7	CO521425	3530_1.14
751	14.8	82.2	663	5	BO064289	Fgr_4_H15	824	14.8	82.2	704	9	CL828295	OR_CBA004
752	14.8	82.2	663	6	CA385174	666023 NC	825	14.8	82.2	705	5	BU066073	Fgr_9_ET12
753	14.8	82.2	663	6	CD208359	H51_37_C1	826	14.8	82.2	705	6	CF006594	QBI20H09.
754	14.8	82.2	664	8	BE667693	CH240_224	827	14.8	82.2	706	5	EX912416	EX912416

Mon Dec 6 12:24:35 2004

us-10-050-189a-6.rst

Page 7

C 828	14.8	82.2	706	7	CO017074	EST87456	C 901	14.8	82.2	757	8	BH715281	BOMHD04TR
C 829	14.8	82.2	706	8	BH201543		C 902	14.8	82.2	759	6	CF002801	CF002801
C 830	14.8	82.2	707	4	BH750230	Fg02_06B1	C 903	14.8	82.2	759	7	CN189784	CN189784
C 831	14.8	82.2	707	4	BM347875	BM347875	C 904	14.8	82.2	760	5	BM873776	BM873776
C 832	14.8	82.2	707	4	CR363121	CR363121	C 905	14.8	82.2	760	8	B2532923	B2532923
C 833	14.8	82.2	708	4	BG836635	BG836635	C 906	14.8	82.2	761	5	BX881221	BX881221
C 834	14.8	82.2	708	4	BG837543	BG837543	C 907	14.8	82.2	761	9	CL133460	CL133460
C 835	14.8	82.2	708	8	AZ263530	AZ263530	C 908	14.8	82.2	762	8	AZ555546	AZ555546
C 836	14.8	82.2	709	8	CK767368	CK767368	C 909	14.8	82.2	763	5	BF458461	BF458461
C 837	14.8	82.2	710	5	BU065685	BU065685	C 910	14.8	82.2	765	2	BX865225	BX865225
C 838	14.8	82.2	710	7	CN185484	CN185484	C 911	14.8	82.2	767	7	CN790929	CN790929
C 839	14.8	82.2	710	7	CO058289	CO058289	C 912	14.8	82.2	768	7	CF672728	CF672728
C 840	14.8	82.2	711	6	CA282671	CA282671	C 913	14.8	82.2	768	9	CG322623	CG322623
C 841	14.8	82.2	711	6	CB377449	CB377449	C 914	14.8	82.2	769	9	CG247559	CG247559
C 842	14.8	82.2	712	5	BU337917	BU337917	C 915	14.8	82.2	770	5	BX873545	BX873545
C 843	14.8	82.2	712	6	CA090040	CA090040	C 916	14.8	82.2	770	9	CG685409	CG685409
C 844	14.8	82.2	712	6	BZ740009	BZ740009	C 917	14.8	82.2	772	6	CF012865	CF012865
C 845	14.8	82.2	713	6	CD209330	CD209330	C 918	14.8	82.2	772	7	CF815840	CF815840
C 846	14.8	82.2	713	8	BH944706	BH944706	C 919	14.8	82.2	772	7	CO214658	CO214658
C 847	14.8	82.2	714	4	CK964734	CK964734	C 920	14.8	82.2	775	6	CD574872	CD574872
C 848	14.8	82.2	717	4	CA113575	CA113575	C 921	14.8	82.2	775	7	CK934101	CK934101
C 849	14.8	82.2	718	6	BH705883	BH705883	C 922	14.8	82.2	779	2	BE034812	BE034812
C 850	14.8	82.2	719	7	CO528582	CO528582	C 923	14.8	82.2	782	7	CO520065	CO520065
C 851	14.8	82.2	719	8	BK305285	BK305285	C 924	14.8	82.2	782	9	CG036391	CG036391
C 852	14.8	82.2	720	5	BK305285	BK305285	C 925	14.8	82.2	784	2	BE542040	BE542040
C 853	14.8	82.2	720	5	BS123278	BS123278	C 926	14.8	82.2	785	5	BX869714	BX869714
C 854	14.8	82.2	721	5	BU063387	BU063387	C 927	14.8	82.2	787	5	BH034402	BH034402
C 855	14.8	82.2	721	9	CG593921	CG593921	C 928	14.8	82.2	789	6	CF006280	CF006280
C 856	14.8	82.2	722	6	CB445384	CB445384	C 929	14.8	82.2	789	7	CO034163	CO034163
C 857	14.8	82.2	723	6	CD423039	CD423039	C 930	14.8	82.2	792	9	CEB19919	CEB19919
C 858	14.8	82.2	724	6	CC655188	CC655188	C 931	14.8	82.2	793	7	CF819290	CF819290
C 859	14.8	82.2	725	4	BC781000	BC781000	C 932	14.8	82.2	794	4	CF813562	CF813562
C 860	14.8	82.2	725	7	CN148302	CN148302	C 933	14.8	82.2	795	6	CF006595	CF006595
C 861	14.8	82.2	726	5	BX880015	BX880015	C 934	14.8	82.2	795	7	CO034549	CO034549
C 862	14.8	82.2	726	6	CA281384	CA281384	C 935	14.8	82.2	795	9	CG415224	CG415224
C 863	14.8	82.2	727	4	BM348044	BM348044	C 936	14.8	82.2	798	9	CO005244	CO005244
C 864	14.8	82.2	727	4	CF665159	CF665159	C 937	14.8	82.2	798	9	CL834970	CL834970
C 865	14.8	82.2	727	8	BH430289	BH430289	C 938	14.8	82.2	799	9	CG211185	CG211185
C 866	14.8	82.2	727	8	BG836351	BG836351	C 939	14.8	82.2	803	3	CF668265	CF668265
C 867	14.8	82.2	730	4	CA135086	CA135086	C 940	14.8	82.2	806	8	BZ811270	BZ811270
C 868	14.8	82.2	730	6	CO123836	CO123836	C 941	14.8	82.2	807	5	BP141963	BP141963
C 869	14.8	82.2	731	6	CA276027	CA276027	C 942	14.8	82.2	807	9	CNS04501	CNS04501
C 870	14.8	82.2	733	6	CN812805	CN812805	C 943	14.8	82.2	808	9	CR528586	CR528586
C 871	14.8	82.2	733	4	BG363115	BG363115	C 944	14.8	82.2	810	7	CB567047	CB567047
C 872	14.8	82.2	734	4	AJ598744	AJ598744	C 945	14.8	82.2	812	6	BM824242	BM824242
C 873	14.8	82.2	735	9	BM338375	BM338375	C 946	14.8	82.2	814	5	BX882424	BX882424
C 874	14.8	82.2	736	4	CA254872	CA254872	C 947	14.8	82.2	814	9	CG607061	CG607061
C 875	14.8	82.2	737	6	CO020561	CO020561	C 948	14.8	82.2	815	3	AY104121	AY104121
C 876	14.8	82.2	737	7	CG259913	CG259913	C 949	14.8	82.2	815	3	AA942156	AA942156
C 877	14.8	82.2	738	6	CA210883	CA210883	C 950	14.8	82.2	818	1	CN253574	CN253574
C 878	14.8	82.2	738	6	CN148444	CN148444	C 951	14.8	82.2	821	9	CG108294	CG108294
C 879	14.8	82.2	738	8	BZ540750	BZ540750	C 952	14.8	82.2	821	9	CG108294	CG108294
C 880	14.8	82.2	739	4	BG836580	BG836580	C 953	14.8	82.2	827	7	CNS02147	CNS02147
C 881	14.8	82.2	739	8	CA247700	CA247700	C 954	14.8	82.2	827	9	CO016429	CO016429
C 882	14.8	82.2	741	6	CA243297	CA243297	C 955	14.8	82.2	828	7	CO082481	CO082481
C 883	14.8	82.2	743	6	CF013380	CF013380	C 956	14.8	82.2	829	7	CF817032	CF817032
C 884	14.8	82.2	743	6	BZ519324	BZ519324	C 957	14.8	82.2	831	7	CO006742	CO006742
C 885	14.8	82.2	744	5	CF012864	CF012864	C 958	14.8	82.2	833	7	CN204728	CN204728
C 886	14.8	82.2	745	5	CN125686	CN125686	C 959	14.8	82.2	833	8	BZ679226	BZ679226
C 887	14.8	82.2	745	7	BZ199901	BZ199901	C 960	14.8	82.2	834	6	CB622839	CB622839
C 888	14.8	82.2	745	8	BZ677522	BZ677522	C 961	14.8	82.2	834	6	BZ142349	BZ142349
C 889	14.8	82.2	746	8	CG367980	CG367980	C 962	14.8	82.2	835	7	CO033630	CO033630
C 890	14.8	82.2	746	9	BZ519324	BZ519324	C 963	14.8	82.2	838	7	CK200133	CK200133
C 891	14.8	82.2	747	6	BZ519324	BZ519324	C 964	14.8	82.2	838	7	CO030131	CO030131
C 892	14.8	82.2	747	6	CA360226	CA360226	C 965	14.8	82.2	840	7	CN316364	CN316364
C 893	14.8	82.2	750	6	CF007121	CF007121	C 966	14.8	82.2	842	9	CL209971	CL209971
C 894	14.8	82.2	750	1	AV928554	AV928554	C 967	14.8	82.2	847	9	CR528585	CR528585
C 895	14.8	82.2	751	1	BX871789	BX871789	C 968	14.8	82.2	853	6	CA128825	CA128825
C 896	14.8	82.2	752	5	CF006454	CF006454	C 969	14.8	82.2	853	6	CG259922	CG259922
C 897	14.8	82.2	754	6	CF602737	CF602737	C 970	14.8	82.2	854	6	CA134468	CA134468
C 898	14.8	82.2	754	9	CL265932	CL265932	C 971	14.8	82.2	855	6	CNS07944	CNS07944
C 899	14.8	82.2	754	9	CL265932	CL265932	C 972	14.8	82.2	857	9		
C 900	14.8	82.2	754	9	CL265932	CL265932	C 973	14.8	82.2	857	9		

74 14.8 82.2 859 7 CFB24787 CFB24787 EST702169
C 975 14.8 82.2 863 7 C0030855 C0030855 EST809239
C 976 14.8 82.2 864 7 C0014376 C0014376 EST784758
C 977 14.8 82.2 872 9 C0263606 C0263606 G0MUK377H
C 978 14.8 82.2 873 2 BR303626 BR303626 601886265
C 979 14.8 82.2 873 7 CFB19514 CFB19514 EST696896
C 980 14.8 82.2 873 7 C1687900 C1687900 PRI0148A
C 981 14.8 82.2 875 9 C0302499 C0302499 C03CB53TV
C 982 14.8 82.2 876 7 C0008648 C0008648 EST7965983
C 983 14.8 82.2 876 7 C0008648 C0008648 EST7965983
C 984 14.8 82.2 877 7 C0005527 C0005527 EST793862
C 985 14.8 82.2 878 3 AY110987 AY110987 Zee may
C 986 14.8 82.2 881 8 A2128036 A2128036 CSJNB008
C 987 14.8 82.2 882 7 C0010892 C0010892 EST799227
C 988 14.8 82.2 887 2 BBS56982 BBS56982 601332443
C 989 14.8 82.2 889 9 C0377810 C0377810 C03DV77TH
C 990 14.8 82.2 893 6 CD458627 CD458627 FG08.1212
C 991 14.8 82.2 895 9 C0035407 C0035407 PUFA12TD
C 992 14.8 82.2 895 9 C0383781 C0383781 CG2C117TH
C 993 14.8 82.2 896 1 A1066632 A1066632 mgsa0003C
C 994 14.8 82.2 896 6 C0662384 C0662384 UCSHY8.0
C 995 14.8 82.2 896 7 C0025150 C0025150 EST80534
C 996 14.8 82.2 898 1 A1666471 A1666471 AL666471
C 997 14.8 82.2 900 5 B0618940 B0618940 RNOSEQ1F0
C 998 14.8 82.2 900 5 B0618946 B0618946 RNOSEQ1F0
C 999 14.8 82.2 900 5 B0619165 B0619165 RNOSEQ4E0
1000 14.8 82.2 900 5 B0619196 B0619196 RNOSEQ4H0

ALIGNMENTS

RESULT 1 164 bp DNA linear GSS 02-OCT-2003
LOCUS CG665075
DEFINITION OST152977 Mus musculus 1295v/Ev Mus musculus genomic clone
ACCESSION CG665075
VERSION CG665075.1 GI:37488924
KEYWORDS GSS.

SOURCE

ORGANISM

Mus musculus (house mouse)
Mus musculus
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Scurionathi; Muridae; Mus.
1 (bases 1 to 164)

REFERENCE

AUTHORS

Zambrowicz, B.P., Abuin, A., Ramirez-Solis, R., Richter, L.J.,
Piggett, J., BeltrandelRio, H., Buxton, E.C., Edwards, J., Finch, R.A.,
Fridde, C.J., Gupta, A., Hansen, G., Hu, Y., Huang, W., Jaing, C.,
Key, B.W., Jr., Kipp, P., Kohauf, B., Ma, Z.-Q., Markesich, D.,
Payne, R., Potter, D.G., Qian, N., Shaw, J., Schrick, J., Shi, Z.-Z.,
Sparks, M.U., Van Sligtenhorst, I., Vogel, P., Walke, W., Xu, N.,
Zhu, Q., Person, C., and Sands, A.T.
Mnk1 kinase deficiency lowers blood pressure in mice: a gene-trap
screen to identify potential targets for therapeutic intervention
Proc. Natl. Acad. Sci. U.S.A. 100 (24), 14109-14114 (2003)

TITLE

Mnk1 kinase deficiency lowers blood pressure in mice: a gene-trap
screen to identify potential targets for therapeutic intervention
Proc. Natl. Acad. Sci. U.S.A. 100 (24), 14109-14114 (2003)

COMMENT

OmniBank
Lexicon Genetics Incorporated
4000 Research Forest Drive, The Woodlands, TX 77381, USA
Email: materials@lexgen.com
Gene trap sequence tag generated by 3' RACE from mouse ES cells as
described in Zambrowicz et al (Nature. 1998 Apr 9;392(6676):608-11)
Class: Gene Trap.

FEATURES

source

1..164
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="1295v/Ev"
/db_xref="taxon:10090"
/clone="OST152977"
/cell_type="embryonic stem cell"
/clone_1ib="Mus musculus 1295v/Ev"

ORIGIN

Query Match 94.4%; Score 17; DB 9; Length 164;
Best Local Similarity 100.0%; Pred. No. 8e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY

1 GAGACACACAGATTG 17
|||||

14 GAGACACACAGATTG 30

RESULT 2

CNS59318 383 bp DNA linear GSS 01-OCT-2003
LOCUS CG559318
DEFINITION OST178045 Mus musculus 1295v/Ev Mus musculus genomic clone
ACCESSION CG559318
VERSION CG559318.1 GI:37345905
KEYWORDS GSS.

SOURCE

ORGANISM

Mus musculus (house mouse)
Mus musculus
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Scurionathi; Muridae; Mus.
1 (bases 1 to 383)

REFERENCE

AUTHORS

Zambrowicz, B.P., Abuin, A., Ramirez-Solis, R., Richter, L.J.,
Piggett, J., BeltrandelRio, H., Buxton, E.C., Edwards, J., Finch, R.A.,
Fridde, C.J., Gupta, A., Hansen, G., Hu, Y., Huang, W., Jaing, C.,
Key, B.W., Jr., Kipp, P., Kohauf, B., Ma, Z.-Q., Markesich, D.,
Payne, R., Potter, D.G., Qian, N., Shaw, J., Schrick, J., Shi, Z.-Z.,
Sparks, M.U., Van Sligtenhorst, I., Vogel, P., Walke, W., Xu, N.,
Zhu, Q., Person, C., and Sands, A.T.
Mnk1 kinase deficiency lowers blood pressure in mice: a gene-trap
screen to identify potential targets for therapeutic intervention
Proc. Natl. Acad. Sci. U.S.A. 100 (24), 14109-14114 (2003)
Contract: Zambrowicz BP

COMMENT

TITLE

Lexicon Genetics Incorporated
4000 Research Forest Drive, The Woodlands, TX 77381, USA
Email: materials@lexgen.com
Gene trap sequence tag generated by 3' RACE from mouse ES cells as
described in Zambrowicz et al (Nature. 1998 Apr 9;392(6676):608-11)
Class: Gene Trap.

JOURNAL

ORIGIN

OmniBank
Lexicon Genetics Incorporated
4000 Research Forest Drive, The Woodlands, TX 77381, USA
Email: materials@lexgen.com
Gene trap sequence tag generated by 3' RACE from mouse ES cells as
described in Zambrowicz et al (Nature. 1998 Apr 9;392(6676):608-11)
Class: Gene Trap.

FEATURES

source

1..383
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="1295v/Ev"
/db_xref="taxon:10090"
/clone="OST178045"
/cell_type="embryonic stem cell"
/clone_1ib="Mus musculus 1295v/Ev"

ORIGIN

Query Match 94.4%; Score 17; DB 9; Length 383;
Best Local Similarity 100.0%; Pred. No. 8.6e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY

1 GAGACACACAGATTG 17
|||||

14 GAGACACACAGATTG 30

Db

14 GAGACACACAGATTG 30

RESULT 3

CNS07414 936 bp DNA linear GSS 07-JUL-2001
LOCUS BA0AB024E03 of library BA0AB from strain CLB 210 of
DEFINITION Kluveromyces lactic, genomic survey sequence.
ACCESSION AL428222
VERSION AL428222.1 GI:12211416
KEYWORDS GSS.

SOURCE

ORGANISM

Kluveromyces lactic
Kluveromyces lactic
Bukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
Saccharomycetales; Saccharomycetaceae; Kluveromyces.

REFERENCE 1 (bases 1 to 936)
 AUTHORS Souciat,J.L., Aigle,M., Artiguenave,F., Blandin,G., Bojotin-Fukuhara,M., Bon,E., Broctier,P., Casaregola,S., deMontigny,J., Dujon,B., Durren,P., Lepingle,A., Llorente,B., Malpertuy,A., Neuvéglise,C., Ozler-Kalogiroupolos,O., Pottier,S., Sarrin,M., Tekala,F., Toffano-Nioche,C., Wesolowski-Louvel,M., Wincker,P., and Weissenbach,J.
 TITLE Genomic exploration of the hemiascomycetous yeasts: 1. A set of yeast species for molecular evolution studies
 JOURNAL FEMS Lett. 487 (1), 3-12 (2000)
 MEDLINE 20584711
 PUBMED 11152876
 REFERENCE 2 (bases 1 to 936)
 AUTHORS Bojotin-Fukuhara,M., Toffano-Nioche,C., Artiguenave,F., Duchateau-Nguyen,G., Lemaire,M., Marmesisse,R., Montrocher,R., Robert,C., Termier,M., Wincker,P., and Wesolowski-Louvel,M.
 TITLE Genomic exploration of the hemiascomycetous yeasts: 11. Kluyveromyces lactis
 JOURNAL FEMS Lett. 487 (1), 66-70 (2000)
 MEDLINE 20584721
 PUBMED 11152886
 REFERENCE 3 (bases 1 to 936)
 AUTHORS Direct Submission
 TITLE Submitted (08-SEP-2000) Genoscope - Centre National de Sequencage, 2 rue Gaston Cremieux, CP 5706, 91057 EVRY cedex, FRANCE. (E-mail: sequef@genoscope.cns.fr - Web : www.genoscope.cns.fr)
 JOURNAL This GSS is part of a random genomic sequencing program of thirteen yeast species: Saccharomyces bayanus var. uvarum, Saccharomyces exiguus, Saccharomyces servazii, Zygosaccharomyces rouxii, Saccharomyces kluyveri, Kluyveromyces thermotolerans, Kluyveromyces lactis var. lactis, Kluyveromyces marxianus var. marxianus, Pichia angusta, Debaryomyces hansenii var. hansenii, Pichia sorbitophila, Candida tropicalis and Yarrowia lipolytica. Genomic inserts of 3 to 5 kb were prepared and both extremities were sequenced. See keywords for description of this sequence and for the sequence of the other extremity of this insert.
 FEATURES
 source 1..936
 /organism="Kluyveromyces lactis"
 /mol_type="genomic DNA"
 /strain="CUIB 210"
 /variety="lactis"
 /db_xref="taxon:28985"
 /clone="BA0AB024E03"
 /clone_11b="BA0AB"
 complement(463..>936)
 /note="Similar to Saccharomyces cerevisiae ORF YDL091c [weak similarity to mouse PAF1 protein]"
 evidence=not_experimental

ORIGIN
 Query Match 94.4%; Score 17; DB 9; Length 936;
 Best Local Similarity 100.0%; Pred. No. 9.3e+02;
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 1 GAGAACACACAGATTGCG 17
 |||||
 875 GAGAACACACAGATTGCG 859

RESULT 4
 LOCUS BG410386 179 bp mRNA linear EST 13-MAR-2001
 DEFINITION 947045H08.xl 947 - 2 week shoot from Barkan lab Zea mays cDNA, mRNA
 SEQUENCE.
 ACCESSION BG410386
 KEYWORDS EST.
 SOURCE Zea mays
 ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD

REFERENCE 1 (bases 1 to 179)
 AUTHORS Walbot,V.
 TITLE Maize ESTs from various cDNA libraries sequenced at Stanford University
 JOURNAL Unpublished (1999)
 COMMENT Department of Biological Sciences
 Stanford University
 855 California Ave, Palo Alto, CA 94304, USA
 Tel: 650 723 2227
 Fax: 650 725 8221
 Email: walbot@stanford.edu
 Plate: 947045 row: H column: 08.
 FEATURES
 source 1..179
 /organism="Zea mays"
 /mol_type="mRNA"
 /cultivar="B73"
 /db_xref="taxon:4577"
 /tissue_type="leaf and stem, including leaf base"
 /dev_stage="2 week old seedling (3 leaves)"
 /lab_host="Xl1-Blue"
 /clone_11b="947 - 2 week shoot from Barkan lab"
 /note="Organ: shoot; Vector: Lambda ZAP (pBluescript SK-); Site 1: EcoRI; Site 2: XhoI; Directionally cloned using Stratagene's Unizap XR cDNA cloning kit with the 5' end at the EcoRI site. The library represents 8 x 10⁵ independent recombinant phage. The plants were greenhouse grown."
 ORIGIN
 Query Match 91.1%; Score 16.4; DB 4; Length 179;
 Best Local Similarity 94.4%; Pred. No. 1.6e+03;
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

DB 1 GAGAACACACAGATTGCG 18
 |||||
 34 GAGAACACACAGATTGCG 17

RESULT 5
 LOCUS CD460790 357 bp mRNA linear EST 14-JUN-2004
 DEFINITION Fg09_01n04 A Fg09_AAFRC BCORC Fusarium graminearum simple_substrate
 Gibberella zeae cDNA clone Fg09_01n04, mRNA sequence.
 ACCESSION CD460790
 VERSION CD460790.2 GI:48688810
 KEYWORDS EST.
 SOURCE Gibberella zeae
 ORGANISM Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes; Hypocreomycetidae; Hypocreales; Nectriaceae; Gibberella.
 REFERENCE 1 (bases 1 to 357)
 AUTHORS Watson,R.J., Heyes,R., Couroux,P., De Moors,A., Harris,L.J., Hattori,J.A., Lacroix,C., Masotti,M., Ouellet,T., Robert,L.S., Singh,U.A., Sprott,D.C., and Tinker,N.A.
 TITLE A cDNA library prepared from Fusarium graminearum grown on a simple substrate
 JOURNAL Unpublished (2003)
 COMMENT On Jun 3, 2003 this sequence version replaced gi:31375530.
 CONTACT: Watson, Robert J.
 Eastern Cereal and Oilseed Research Centre
 Agriculture and Agri-Food Canada
 Bldg. 20, Central Experimental Farm, Ottawa, Ontario, K1A 0C6, CANADA
 Tel: (613) 759-1655
 Fax: (613) 759-1701
 Email: watsonr@agr.gc.ca
 FEATURES
 source 1..357
 /organism="Gibberella zeae"
 /mol_type="mRNA"

/strain="DAOM 180378"
 /db_xref="taxon:5518"
 /clone="Fg09_01n04"
 /tissue_type="Mycelium"
 /dev_stage="Asexual"
 /lab_host="E. coli DH10B"
 /clone_lib="Fg09_AAFc_ECORC_Fusarium graminearum simple_substrata"
 /note="Vector: pBluescript II+, site 1: EcoRI; site 2: XhoI; Fusarium graminearum grown on a simple substrate--minimal media supplemented with amino acids."

ORIGIN

Query Match 91.1%; Score 16.4; DB 6; Length 357;
 Best Local Similarity 94.4%; Pred. No. 1.7e+03;
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 1 GAGAACACACAGATTGCG 18
 |||||
 Db 223 GAGAACACACAGATTGCG 240

RESULT 6

AQ210247

LOCUS HS_3115_B2_C12_T7_C1T Approved Human Genomic Sperm Library D Homo
 DEFINITION sapiens genomic clone Plate=3115 Col=24 Row=F, genomic survey
 sequence.

ACCESSION AQ210247.1 GI:3619216
 VERSION AQ210247.1
 KEYWORDS GSS.
 SOURCE Homo sapiens (human)

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 1 (bases 1 to 481)
 Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T., Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and Hood,L.

TITLE Sequence-tagged connectors: A sequence approach to mapping and scanning the human genome
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
 MEDLINE 99380589
 PUBMED 10449764
 COMMENT Contact: Mahairas GG, Wallace JC, Hood L
 High Throughput Sequencing Center
 University of Washington
 401 Queen Anne Avenue North, Seattle, WA 98109, USA
 Tel: (206) 616-3618
 Fax: (206) 616-3887
 Email: jwallace@u.washington.edu
 Sequence Tagged Connector
 Plate: 3115 row: F column: 24
 Class: BAC ends
 High quality sequence stop: 481.

FEATURES

source

1..481
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /clone="Plate=3115 Col=24 Row=F"
 /sex="male"
 /note="CIT Approved Human Genomic Sperm Library D"
 /note="Organ: sperm; Vector: pBluescript II; BAC Clones in E-Coli DH10B"

ORIGIN

Query Match 91.1%; Score 16.4; DB 8; Length 481;
 Best Local Similarity 94.4%; Pred. No. 1.8e+03;
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 1 GAGAACACACAGATTGCG 18
 |||||
 |||||

Db

242 GAGAACACACAGATTGCG 255

RESULT 7

CD657925/c

LOCUS ECESTE5F909.Y1 Eimeria tenella M5-6 Excised cDNA Eimeria tenella
 DEFINITION CDNA 5', mRNA sequence.

ACCESSION CD657925
 VERSION CD657925.1 GI:3190054
 KEYWORDS EST.
 SOURCE Eimeria tenella
 ORGANISM Eimeria tenella
 Eukaryota; Alveolata; Apicomplexa; Coccidia; Eimeriida; Eimeriidae; Eimeria.

REFERENCE

1 (bases 1 to 504)
 Liberator,P., Diaz,C., Tang,K., Marra,M., Hillier,L., Kucaba,T., Martin,J., Wylie,T., Underwood,K., Stepec,M., Theising,B., Allen,M., Bowers,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R., Ritter,E., Kohn,S., Florence,N., Shin,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R., Wilson,R. and Sibley,D.
 WashU-Merck Eimeria tenella project
 Unpublished (1999)
 Contact: David Sibley, Ph.D.
 WashU-Merck Eimeria tenella project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@watson.wustl.edu
 Contact David Sibley (toxest@borcim.wustl.edu) for further information relating to organism, libraries, or clone availability.
 Seq primer: -40RP from Gibco.

FEATURES

source

1..504
 /organism="Eimeria tenella"
 /mol_type="mRNA"
 /db_xref="taxon:5802"
 /dev_stage="IS:18"
 /lab_host="E. coli DH10B (GeneHog, Invitrogen, Inc)"
 /clone_lib="Eimeria tenella M5-6 Excised cDNA"
 /note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2: XhoI; Sporozoites were obtained from in vitro sporulated and excysted oocysts of E. tenella grown in chickens. cDNA was synthesized from poly RNA using an oligo-dT primer containing a XhoI site. Following second strand synthesis, EcoRI adapters were ligated to the cDNA and products were size-selected on sephacryl S500. The cDNA were ligated to EcoRI/XhoI prepared lambda ZapII (Stratagene). The primary library was mass excised using Exsist helper phage (Stratagene). The phagemids were precipitated with PEG 8000, extracted with phenol/chloroform and electroporated into DH10B cells. The library may contain a small percentage of host or bacterial contaminants. Library materials provided by: Paul Liberator, Merck Research Labs Library constructed by: Paul Liberator, Merck Research Labs."

ORIGIN

Query Match 91.1%; Score 16.4; DB 6; Length 504;
 Best Local Similarity 94.4%; Pred. No. 1.8e+03;
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 1 GAGAACACACAGATTGCG 18
 |||||
 Db 347 GAGAACACACAGATTGCG 330

RESULT 8
 LOCUS BM076070/c
 DEFINITION BM076070 514 bp mRNA linear EST 13-NOV-2001
 MEST364-D04.T3 ISUM5-RN Zea mays cDNA clone MEST364-D04 3', mRNA

RESULT 10
 BI823206/c
 LOCUS
 DEFINITION
 BI823206 1138 bp mRNA linear EST 04-OCT-2001
 603039469F1 NIH_MGC_115 Homo sapiens cDNA clone IMAGE:5180272 5',
 mRNA sequence.
 ACCESSION
 BI823206
 VERSION
 BI823206.1 GI:15934756
 KEYWORDS
 EST.
 SOURCE
 Homo sapiens (human)
 ORGANISM
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE
 1 (bases 1 to 1138)
 NIH-MGC <http://mgc.nci.nih.gov/>.
 TITLE
 National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL
 UNPUBLISHED (1999)
 COMMENT
 Contact: Robert Strausberg, Ph.D.
 Email: rgs@nih.gov
 Tissue Procurement: Life Technologies, Inc.
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LIML)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LIML at:
<http://image.llnl.gov>
 Plate: LIML149 row: j column: 17
 High quality sequence stop: 324.
 Location/Qualifiers
 1..1138
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5180272"
 /lab_host="DH10B"
 /clone_id="NIH_MGC_115"
 /note="Organ: pooled brain, lung, testis; Vector:
 PCW-SPORT6; Site 1: NotI; Site 2: EcoRV (destroyed); RNA
 source anonymous pool of 6 male brains, age range 23-27; 1
 male lung, age 27; and 1 male testis, age 69. Library is
 oligo-dT primed and directionally cloned (EcoRV site is
 destroyed upon cloning). Average insert size 1.8 kb,
 insert size range 1-3 kb. Library is normalized and
 enriched for full-length clones and was constructed by C.
 Gruber (Invitrogen). Research Genetics tracking code
 021. Note: this is a NIH_MGC Library."
 ORIGIN
 Query Match 91.1%; Score 16.4; DB 4; Length 1138;
 Best Local Similarity 94.4%; Pred. No. 1.9e+03;
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 GAGAACAAAGATTGCG 18
 |||||
 DB 1011 GAGAACAAAGATTGCG 994
 |||||
 RESULT 11
 BE574801
 LOCUS
 DEFINITION
 BE574801 368 bp mRNA linear EST 07-DEC-2001
 H19 Triphysaria versicolor root-tip, early DMBQ-induced transcript
 cDNA library Triphysaria versicolor cDNA, mRNA sequence.
 ACCESSION
 BE574801
 VERSION
 BE574801.1 GI:12001131
 KEYWORDS
 EST.
 SOURCE
 Triphysaria versicolor
 ORGANISM
 Triphysaria versicolor
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta, eudicotyledons; core eudicots;
 asterids; Lamiales; Lamiaceae; Orobanchaceae; Rhinanthaceae;
 Triphysaria.
 REFERENCE
 1 (bases 1 to 368)

AUTHORS
 TITLE
 JOURNAL
 MEDLINE
 PUBMED
 COMMENT
 Contact: Yoder, J.I.
 John I. Yoder Research Lab, Dept. of Vegetable Crops
 University of California at Davis
 137 Asmunsen Hall, One Shields Drive, Davis, CA 95616, USA
 Tel: 530 752 1741
 Fax: 530 752 9659
 Email: jyoder@ucdavis.edu
 length = 368 bp.
 Location/Qualifiers
 1..368
 /organism="Triphysaria versicolor"
 /mol_type="mRNA"
 /db_xref="taxon:64093"
 /tissue_type="root-tips ~5mm length"
 /dev_stage="3-4 weeks growth"
 /lab_host="E. coli"
 /clone_id="Triphysaria versicolor root-tip, early
 DMBQ-induced transcript cDNA library"
 /note="Vector: PCR2.1 TA Cloning System, Invitrogen,
 Carlsbad, CA, PCR-based suppression subtractive
 hybridization cDNA library"
 ORIGIN
 Query Match 88.9%; Score 16; DB 2; Length 368;
 Best Local Similarity 100.0%; Pred. No. 2.7e+03;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 3 GACACAAAGATTGCG 18
 |||||
 DB 86 GACACAAAGATTGCG 101
 |||||
 RESULT 12
 CE085246
 LOCUS
 DEFINITION
 CE085246 426 bp DNA linear GSS 24-SEP-2003
 tigr-gss-dog-17000359005989 Dog Library Canis familiaris genomic,
 genomic survey sequence.
 ACCESSION
 CE085246
 VERSION
 CE085246.1 GI:35152092
 KEYWORDS
 GSS.
 SOURCE
 Canis familiaris (dog)
 ORGANISM
 Canis familiaris
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
 REFERENCE
 1 (bases 1 to 426)
 Kirkness, E.F., Bafna, V., Halpern, A.L., Levy, S., Remington, K.,
 Rusch, D.B., Delcher, A.L., Pop, M., Wang, W., Fraser, C.M. and
 Venter, J.C.
 The dog genome: survey sequencing and comparative analysis
 Science 301 (5641), 1898-1903 (2003)
 MEDLINE
 PUBMED
 COMMENT
 Contact: Kirkness EF
 The Institute for Genomic Research
 Department of Eukaryotic Genomics, TIGR, 9712 Medical Center Drive,
 Rockville, MD 20850, USA
 Tel: 301-838-0200
 Fax: 301-838-0208
 Email: ekirkness@tigr.org
 Class: shotgun.
 Location/Qualifiers
 1..426
 /organism="Canis familiaris"
 /mol_type="genomic DNA"
 /strain="Standard Poodle"
 /db_xref="taxon:9615"
 /clone_id="Dog Library"

ORIGIN /note="Site 1: Betxi, Libraries were prepared from peripheral blood"

Query Match 88.9%; Score 16; DB 9; Length 426;
Best Local Similarity 100.0%; Pred. No. 2.9e+03;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAGACACACAGATTC 16
|||||
DB 111 GAGACACACAGATTC 126

RESULT 13 CK130476 646 bp mRNA linear EST 02-DEC-2003
LOCUS AT081813prtime AT Drosophila melanogaster adult testes POTB7
DEFINITION Drosophila melanogaster cDNA clone AT081813 similar to CG9905:
Fham0009905 GO: [] located on: 4 102D1-102D1; 08/14/2002, mRNA
sequence.

ACCESSION CK130476
VERSION CK130476
KEYWORDS GI:38622412
SOURCE EST.
ORGANISM Drosophila melanogaster (fruit fly)
Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.

REFERENCE 1 (bases 1 to 646)
Stapleton, M., Brokstein, P., Hong, L., Agbayan, A., Baxter, E.,
Berman, B., Carlson, J., Champe, M., Chavez, C., Chew, M., Dorsett, V.,
Farfan, D., Frise, E., George, R., Gonzalez, M., Guatin, H., Harris, N.,
Li, P., Liao, G., Miranda, A., Mitra, S., Mungall, C. J., Nunoo, J.,
Paele, J., Paragas, V., Park, S., Phouanavong, S., Wan, K., Yu, C.,
Lewis, S. E., Celniker, S. and Rubin, G. M.
BDGP/HMT AT Drosophila EST Project
Unpublished (2000)
Other ESTs: AT08181.5prtime
Contact: Stapleton, M.
BDGP
Lawrence Berkeley National Lab
One Cyclotron Rd, Berkeley, CA 94720, USA
Fax: 510 486 6798
Email: http://www.fruitfly.org/EST_est@fruitfly.berkeley.edu
Based upon the presence of a XhoI site followed by a run of 14 or
more T residues at the beginning of the sequence, this clone was
polyadenylated. The resulting Poly-T sequence has been removed. hit
genomic AF003843: arm:4 [434780,751419] estimated-cyto:102c1-102a1:
02/05/2002

FEATURES
source Plate: AT.81 row: G column: 9
High quality sequence stop: 560.
Location/Qualifiers

1..646
/organism="Drosophila melanogaster"
/mol_type="mRNA"
/db_xref="taxon:7227"
/clone="AT08181"
/sex="male"
/dev_stage="0-3 day old Ore-R males"
/lab_host="Plates AT.10-AT.120; DHS-alpha. Plates
AT.121-AT.319; DHS-alpha TONa"
/clone_lib="AT Drosophila melanogaster adult testes POTB7"
/note="Organ: ADULT testes; Vector: POTB7; Site 1: EcoRI;
Site 2: XhoI; The mRNA for the testis library was made
from testes and seminal vesicles hand dissected from 0-3
day old Ore-R males. RNA kindly provided by the lab of
Margaret Fuller. Sized fractionated cDNAs were directly
ligated into POTB7. Plasmid cDNA library."

ORIGIN

Query Match 88.9%; Score 16; DB 7; Length 646;
Best Local Similarity 100.0%; Pred. No. 2.9e+03;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 AGAACACACAGATTCG 17
|||||
DB 583 AGAACACACAGATTCG 598

RESULT 14 B2178809 690 bp DNA linear GSS 11-OCT-2002
LOCUS CH230-443N20.TVB CHORI-230 Segment 2 Rattus norvegicus genomic
clone CH230-443N20, genomic survey sequence.
B2178809
B2178809.1 GI:23822940

ACCESSION B2178809.1 GI:23822940
VERSION B2178809.1
KEYWORDS GSS.
SOURCE Rattus norvegicus (Norway rat)
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE 1 (bases 1 to 690)
Zhao, S., Shetty, J., Shatsman, S., Tsegaye, G., Geer, K.,
Shvartsbeyn, A., Gebregorgis, B., Overton, L., Russell, D., Chen, D.,
Riggs, F., de Jong, P. and Fraser, C. M.
Rat BAC End Sequences from Library CHORI-230 MboI segment
Unpublished (1999)
Other GSSs: CH230-443N20.TVB
Contact: Shaying Zhao
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: szhao@igr.org
Clones are derived from the rat BAC library CHORI-230
(<http://www.chori.org/bacpac/rat230.htm>). For BAC library
availability, please contact Pieter de Jong (pdejong@mail.cno.org).
Clones may be purchased from BACPAC Resources
(http://www.chori.org/bacpac/or_eirng_information.htm). BAC end
page: http://www.tigr.org/cdb/bac_ends/rat/bac_end_intro.html
Plate: 443 row: N column: 20
Seq primer: 17
Class: BAC ends.
Location/Qualifiers

FEATURES
source 1..690
Location/Qualifiers

/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/strain="BN/SnHsd/MCW"
/db_xref="taxon:10116"
/clone="CH230-443N20"
/sex="Female"
/cell_type="Brain"
/clone_lib="CHORI-230 Segment 2"
/note="Vector: PTARBAC1.3; Site 1: MboI; Site 2: MboI;
CHORI-230 Rat (BN/SnHsd/MCW) BAC library produced by
Pieter de Jong"

ORIGIN

Query Match 88.9%; Score 16; DB 8; Length 690;
Best Local Similarity 100.0%; Pred. No. 2.9e+03;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAGACACACAGATTC 16
|||||
DB 565 GAGACACACAGATTC 580

RESULT 15 CC082521 744 bp DNA linear GSS 16-APR-2003
LOCUS CSU-K33r.3868.T7 CSU-K33r Aedes aegypti genomic clone
DEFINITION CSU-K33r.3868, genomic survey sequence.
ACCESSION CC082521
VERSION CC082521.1 GI:29931623

KEYWORDS GSS.
SOURCE Aedes aegypti (yellow fever mosquito)
ORGANISM Aedes aegypti
Bukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae; Aedes; Stegomyia.
REFERENCE 1 (bases 1 to 744)
AUTHORS Loftus, B., Shetty, J., Severson, D., Brown, S. and Knudson, D.
TITLE End sequencing of Aedes aegypti BACs
JOURNAL Unpublished (2003)
COMMENT Other_GSSs: CSU-K33r.38G8.SP6
Contact: Brendan Loftus
Department of Eukaryotic Genomics
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-3543
Fax: 301-838-0208
Email: ena@tigr.org
Library was provided by Susan Brown and Dennis Knudson at Colorado State University.
Seq primer: T7
Class: BAC ends.

FEATURES
SOURCE location/Qualifiers
1..744
/organism="Aedes aegypti"
/mol_type="genomic DNA"
/strain="Rexville"
/db_xref="taxon:7159"
/clone="CSU-K33r.38G8"
/clone_1db="CSU-K33r"
/note="Vector: pBelBAC11; Site_1: HindIII"

ORIGIN
Query Match 88.9%; Score 16; DB 8; Length 744;
Best Local Similarity 100.0%; Pred. No. 2.9e+03;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 AGAACAACAAGATTTC 17
|||||
Db 544 AGAACAACAAGATTTC 559

RESULT 16
CL491245/c 892 bp DNA linear GSS 01-APR-2004
LOCUS SAIL_554_A05.v2 SAIL Collection Arabidopsis thaliana genomic clone
DEFINITION SAIL_554_A05.v2, genomic survey sequence.
ACCESSION CL491245
VERSION CL491245.1 GI:45975701
KEYWORDS GSS.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eustosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE 1 (bases 1 to 892)
AUTHORS Sessions, A., Burke, E., Presting, G., Aux, G., McElver, J., Paton, D., Dietrich, B., Ho, P., Bacwaden, J., Ko, C., Clarke, J.D., Cotton, D., Bullis, D., Snell, J., Miguel, T., Hutchinson, D., Kimmery, B., Mitzel, T., Katsagiri, F., Glazebrook, J., Law, M. and Goff, S.A.
TITLE A high-throughput Arabidopsis reverse genetics system
JOURNAL Plant Cell 14 (12), 2985-2994 (2002)
MEDLINE 22356987
PUBMED 12468722
COMMENT Contact: Sessions A
Applied Trait Genetics
SynGene Biotechnology Inc.
3054 Cornwalls Rd., Research Triangle Park, NC 27709, USA
Email: allen.sessions@syngenta.com
ABRC Stock Number CS823429; T-DNA left border flanking sequences of Syngenta Arabidopsis Insertion Library (SAIL) lines are available through the Arabidopsis Biological Resource Center (ABRC).
Sequences represent a pool of amplified genomic regions and not

single contiguous sequences.
Class: TDNA tagged.
location/Qualifiers
1..892
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/ecotype="Columbia"
/db_xref="taxon:3702"
/clone="SAIL_554_A05.v2"
/clone_1db="SAIL Collection"
/note="T-DNA left border sequences were isolated using a modified TAIL-PCR strategy"

ORIGIN
Query Match 88.9%; Score 16; DB 9; Length 892;
Best Local Similarity 100.0%; Pred. No. 2.9e+03;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAGACAACAAGATTTC 16
|||||
Db 622 GAGACAACAAGATTTC 607

RESULT 17
BU017470 227 bp mRNA linear EST 23-AUG-2002
LOCUS OHE15P17.YG.ab1 OH_EFGH sunflower RHA260 Helianthus annuus cDNA
DEFINITION clone OHE15P17, mRNA sequence.
ACCESSION BU017470
VERSION BU017470.1 GI:22452990
KEYWORDS EST.
SOURCE Helianthus annuus (common sunflower)
ORGANISM Helianthus annuus
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids; campanulids; Asterales; Asterales; Asteroideae; Heliantheae; Helianthus.
REFERENCE 1 (bases 1 to 227)
AUTHORS Kozik, A., Michelmore, R.W., Knapp, S., Matvienko, M., Rieseberg, L., Lin, H., van Damme, M., Lavelle, D., Chevallier, P., Ziegler, J., Allison, P., Kolman, J., Slabaugh, M.S., Livingston, K., Zhou, Y., Lai, Z., Church, S., Jackson, L. and Bradford, K.
TITLE Lettuce and Sunflower ESTs from the Composite Genome Project
JOURNAL http://compgenome.ucdavis.edu/
COMMENT Unpublished (2002)
CONTACT: Alexander Kozik [R.W.Michelmore]
Department of Vegetable Crops, R.W.Michelmore Lab
University of California at Davis (UCD)
Asmundson Hall, UCD, Davis, CA 95616, USA
Tel: 1-(530)-742-1742
Fax: 1-(530)-752-9659
Email: akozik@ucdavis.edu [michelmore@vegmail.ucdavis.edu]
belongs to contig OH_CA.Contig192, see http://cgpdb.ucdavis.edu/ for details.
Plate: OHE15 row: P column: 17.

FEATURES
SOURCE location/Qualifiers
1..227
/organism="Helianthus annuus"
/mol_type="mRNA"
/cultiivar="RHA260"
/db_xref="taxon:4232"
/clone="OHE15P17"
/lab_host="E.coli"
/clone_1db="OH_EFGH sunflower RHA260"
/note="Vector: pBRCDNA51AB; The library was constructed from 11 different sources of RNA from a single genotype. Separate cDNAs were generated using primers that incorporated unique 5' and 3' tags to distinguish each source of RNA. cDNAs were then pooled, size-fractionated, directionally cloned into a custom medium-copy vector and transformants made with four size classes to minimize size bias. Details of each source of RNA and library construction can be obtained at http://cgpdb.ucdavis.edu/

ORIGIN TAG_TISSUE=shoots environmental stress
TAG_LIB=OH_EFGHJ sunflower RHA280
TAG_SEQ=TCGCAACGGG"

Query Match 85.6%; Score 15.4; DB 5; Length 229;
LOCUS 94.1%; Pred. No. 5.2e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 AGAACACACAGATTCCG 18
|||||
82 AGAAAAACAAGATTCCG 98

RESULT 18 BU018252 229 bp mRNA linear EST 23-AUG-2002
BU018252 OHE18B18.yg.ab1 OH_EFGHJ sunflower RHA280 Helianthus annuus cDNA
LOCUS clone OHE18B18, mRNA sequence.
DEFINITION
ACCESSION BU018252 GI:22453772
VERSION
KEYWORDS
SOURCE Helianthus annuus (common sunflower)
ORGANISM Helianthus annuus
COMMENT Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; campanulids; Asterales; Asteraceae; Asteroideae;
Heliantheae; Helianthus.
1 (bases 1 to 229)
Kozik,A., Michelmore,R.W., Knapp,S., Matvienko,M., Rieseberg,L.,
Lin,H., van Damme,M., Lavelle,D., Chevalier,P., Ziegler,J.,
Ellison,P., Kolkmann,J., Slabaugh,M.S., Livingston,K., Zhou,Y.,
Lai,Z., Church,S., Jackson,L. and Bradford,K.
Lettuce and Sunflower ESTs from the Composite Genome Project
http://compgenome.ucdavis.edu/
Unpublished (2002)
Contact: Alexander Kozik [R.W.Michelmore]
Department of Vegetable Crops, R.W.Michelmore Lab
University of California at Davis (UCD)
Aasmundson Hall, UCD, Davis, CA 95616, USA
Tel: 1-(530)-742-1742
Fax: 1-(530)-752-9659
Email: akozik@ucdavis.edu [michelmoreveg@mail.ucdavis.edu]
belongs to contig OH_CA.Contig192, see http://cgpdb.ucdavis.edu/
for details.
Plate: OHE18 row: B column: 18.

FEATURES
source Location/Qualifiers
1..229
/organism="Helianthus annuus"
/mol_type="mRNA"
/cultiivar="RHA280"
/db_xref="taxon:4232"
/clone="OHE18B18"
/lab_host="E.coli"
/clone_lib="OH_EFGHJ sunflower RHA280"
/note="Vector: pBRCN45flab; The library was constructed
from 11 different sources of RNA from a single genotype.
Separate cDNAs were generated using primers that
incorporated unique 5' and 3' tags to distinguish each
source of RNA. cDNAs were then pooled, size-fractionated,
directionally cloned into a custom medium-copy vector and
transformations made with four size classes to minimize
size bias. Details of each source of RNA and library
construction can be obtained at http://cgpdb.ucdavis.edu/
TAG_SEQ=Not found"

ORIGIN

Query Match 85.6%; Score 15.4; DB 5; Length 229;
Best Local Similarity 94.1%; Pred. No. 5.2e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

2 AGAACACACAGATTCCG 18

Db 81 AGAAAAACAAGATTCCG 97

RESULT 19 BU016050 269 bp mRNA linear EST 23-AUG-2002
BU016050 OHE11P07.yg.ab1 OH_EFGHJ sunflower RHA280 Helianthus annuus cDNA
LOCUS clone OHE11P07, mRNA sequence.
DEFINITION
ACCESSION BU016050 GI:22451570
VERSION
KEYWORDS
SOURCE Helianthus annuus (common sunflower)
ORGANISM Helianthus annuus
COMMENT Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; campanulids; Asterales; Asteraceae; Asteroideae;
Heliantheae; Helianthus.
1 (bases 1 to 269)
Kozik,A., Michelmore,R.W., Knapp,S., Matvienko,M., Rieseberg,L.,
Lin,H., van Damme,M., Lavelle,D., Chevalier,P., Ziegler,J.,
Ellison,P., Kolkmann,J., Slabaugh,M.S., Livingston,K., Zhou,Y.,
Lai,Z., Church,S., Jackson,L. and Bradford,K.
Lettuce and Sunflower ESTs from the Composite Genome Project
http://compgenome.ucdavis.edu/
Unpublished (2002)
Contact: Alexander Kozik [R.W.Michelmore]
Department of Vegetable Crops, R.W.Michelmore Lab
University of California at Davis (UCD)
Aasmundson Hall, UCD, Davis, CA 95616, USA
Tel: 1-(530)-742-1742
Fax: 1-(530)-752-9659
Email: akozik@ucdavis.edu [michelmoreveg@mail.ucdavis.edu]
belongs to contig OH_CA.Contig192, see http://cgpdb.ucdavis.edu/
for details.
Plate: OHE11 row: P column: 07.

FEATURES
source Location/Qualifiers
1..269
/organism="Helianthus annuus"
/mol_type="mRNA"
/cultiivar="RHA280"
/db_xref="taxon:4232"
/clone="OHE11P07"
/lab_host="E.coli"
/clone_lib="OH_EFGHJ sunflower RHA280"
/note="Vector: pBRCN45flab; The library was constructed
from 11 different sources of RNA from a single genotype.
Separate cDNAs were generated using primers that
incorporated unique 5' and 3' tags to distinguish each
source of RNA. cDNAs were then pooled, size-fractionated,
directionally cloned into a custom medium-copy vector and
transformations made with four size classes to minimize
size bias. Details of each source of RNA and library
construction can be obtained at http://cgpdb.ucdavis.edu/
TAG_TISSUE=shoots environmental stress
TAG_LIB=OH_EFGHJ sunflower RHA280
TAG_SEQ=TCGCAACGGG"

ORIGIN

Query Match 85.6%; Score 15.4; DB 5; Length 269;
Best Local Similarity 94.1%; Pred. No. 5.3e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 AGAACACACAGATTCCG 18
|||||
82 AGAAAAACAAGATTCCG 98

Db

RESULT 20 CD944524 270 bp mRNA linear EST 15-JUL-2003
CD944524 BDJ 72 Genetagi Zea mays cDNA, mRNA sequence.
LOCUS
DEFINITION
ACCESSION CD944524 GI:32792288
VERSION

KEYWORDS
SOURCE
ORGANISM
EST.
Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
1 (bases 1 to 270)
Genopiante.
Unpublished (2003)
Contact: Genopiante
Genopiante
93, rue Henri Rochefort 91025 EVRY CEDEX France
Tel: 33 1 69 47 54 00
Fax: 33 1 69 47 54 10
This sequence has been generated in the framework of the french
plant genomics programme 'Genopiante' (<http://www.genopiante.com>
and <http://genopiante-info.infobiogen.fr>).

FEATURES
source
1..270
/organism="Zea mays"
/mol_type="mRNA"
/cultivar="mixture"
/db_xref="taxon:4577"
/clone_1lb="Genetags1"

ORIGIN
Query Match 85.6%; Score 15.4; DB 6; Length 270;
Best Local Similarity 94.1%; Pred. No. 5.3e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 1 GAGAACACACAGATTGC 17
86 GAGAACACACATGATTGC 102

RESULT 21
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
CD946927 270 bp mRNA linear EST 15-JUL-2003
REV 27 Genetags1 Zea mays cDNA, mRNA sequence.
CD946927
CD946927.1 GI:32794691
EST.
Zea mays
Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
1 (bases 1 to 270)
Genopiante.
Unpublished (2003)
Contact: Genopiante
Genopiante
93, rue Henri Rochefort 91025 EVRY CEDEX France
Tel: 33 1 69 47 54 00
Fax: 33 1 69 47 54 10
This sequence has been generated in the framework of the french
plant genomics programme 'Genopiante' (<http://www.genopiante.com>
and <http://genopiante-info.infobiogen.fr>).

FEATURES
source
1..270
/organism="Zea mays"
/mol_type="mRNA"
/cultivar="mixture"
/db_xref="taxon:4577"
/clone_1lb="Genetags1"

ORIGIN
Query Match 85.6%; Score 15.4; DB 6; Length 270;
Best Local Similarity 94.1%; Pred. No. 5.3e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 1 GAGAACACACAGATTGC 17
86 GAGAACACACATGATTGC 102

QY 1 GAGAACACACAGATTGC 17
86 GAGAACACACATGATTGC 102

Db 1 GAGAACACACAGATTGC 17
86 GAGAACACACATGATTGC 102

RESULT 22
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
CD967114 270 bp mRNA linear EST 16-JUL-2003
SES 132 Genetags2 Zea mays cDNA, mRNA sequence.
CD967114
CD967114.1 GI:32827436
EST.
Zea mays
Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
1 (bases 1 to 270)
Genopiante.
Unpublished (2003)
Contact: Genopiante
Genopiante
93, rue Henri Rochefort 91025 EVRY CEDEX France
Tel: 33 1 69 47 54 00
Fax: 33 1 69 47 54 10
This sequence has been generated in the framework of the french
plant genomics programme 'Genopiante' (<http://www.genopiante.com>
and <http://genopiante-info.infobiogen.fr>).

FEATURES
source
1..270
/organism="Zea mays"
/mol_type="mRNA"
/cultivar="mixture"
/db_xref="taxon:4577"
/clone_1lb="Genetags2"

ORIGIN
Query Match 85.6%; Score 15.4; DB 6; Length 270;
Best Local Similarity 94.1%; Pred. No. 5.3e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 1 GAGAACACACAGATTGC 17
86 GAGAACACACATGATTGC 102

RESULT 23
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
CD967248 270 bp mRNA linear EST 16-JUL-2003
SES 61 Genetags2 Zea mays cDNA, mRNA sequence.
CD967248
CD967248.1 GI:32827570
EST.
Zea mays
Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
1 (bases 1 to 270)
Genopiante.
Unpublished (2003)
Contact: Genopiante
Genopiante
93, rue Henri Rochefort 91025 EVRY CEDEX France
Tel: 33 1 69 47 54 00
Fax: 33 1 69 47 54 10
This sequence has been generated in the framework of the french
plant genomics programme 'Genopiante' (<http://www.genopiante.com>
and <http://genopiante-info.infobiogen.fr>).

FEATURES
source
1..270
/organism="Zea mays"
/mol_type="mRNA"
/cultivar="mixture"
/db_xref="taxon:4577"
/clone_1lb="Genetags2"

ORIGIN
Query Match 85.6%; Score 15.4; DB 6; Length 270;
Best Local Similarity 94.1%; Pred. No. 5.3e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 1 GAGAACACACAGATTGC 17
86 GAGAACACACATGATTGC 102

ORIGIN

```

/organism="Zea mays"
/mol_type="mRNA"
/cultivar="mixture"
/db_xref="taxon:4577"
/clone_lib="Genetags2"

```

Query Match 85.6%; Score 15.4; DB 6; Length 270;
 Best Local Similarity 94.1%; Pred. No. 5.3e+03;
 Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GAGAACACAGATTGG 17
 |||||
 Db 185 GAGAACACAGATTGG 169

RESULT 24
 AM977559 273 bp mRNA linear EST 02-JUN-2000
 LOCUS EST389668 MAGR resequences, MAGO Homo sapiens cDNA, mRNA sequence.
 ACCESSION AM977559
 VERSION AM977559.1 GI:8168812
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 273)
 Hegde, P., Qi, R., Abernathy, K., Daarp, S., Gaspard, R., Gay, C.,
 Holt, I.E., Speed, A.T., Sharov, V., Lee, N.H., Yeatman, T.J. and
 Quackenbush, J.
 Assessment of gene expression patterns in a model of colon tumor
 metastasis using a 19,200 element cDNA microarray
 Unpublished (2000)

TITLE The Institute for Genomic Research
 CONTACT John Quackenbush
 JOURNAL 9712 Medical Center Dr., Rockville, MD 20850, USA
 COMMENT Tel: 301 838 3528
 Fax: 301 838 0208
 Email: john@cgisr.org
 Plate: 380

Seq primer: Forward.

FEATURES Location/Qualifiers
 source 1..273
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone_lib="MAGE resequences, MAGO"
 /note="Vector: pBluescriptskm"

ORIGIN

```

Query Match 85.6%; Score 15.4; DB 2; Length 273;  

Best Local Similarity 94.1%; Pred. No. 5.3e+03;  

Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

Qy 2 AGAACACAGATTGG 18
 |||||
 Db 251 AGAACACAGATTGG 267

RESULT 25
 BQ910769 284 bp mRNA linear EST 19-AUG-2002
 LOCUS CHB15C10.yg.ab1 QH ABCDI sunflower RHA601 Helianthus annuus cDNA
 DEFINITION clone QH15C10, mRNA sequence.
 ACCESSION BQ910769
 VERSION BQ910769.1 GI:22309548
 KEYWORDS EST.
 SOURCE Helianthus annuus (common sunflower)
 ORGANISM Helianthus annuus

REFERENCE 1 (bases 1 to 284)
 Kozik, A., Michelmore, R.W., Krapp, S., Matvienko, M., Rieseberg, L.,
 Lin, H., van Damme, M., Lavelle, D., Chevaller, P., Ziegler, J.,
 Ellison, P., Kolman, J., Slabaugh, M.S., Livingston, K., Zhou, Y.,
 Lai, Z., Church, S., Jackson, L. and Bradford, K.
 Lettuce and Sunflower ESTs from the Composite Genome Project
 http://comgenomics.ucdavis.edu/
 Unpublished (2002)
 CONTACT Alexander Kozik [R.W.Michelmore]
 Department of Vegetable Crops, R.W.Michelmore Lab
 University of California at Davis (UCD)
 Aemundson Hall, UCD, Davis, CA 95616, USA
 Tel: 1-(530)-742-1742
 Fax: 1-(530)-752-9659
 Email: akozik@ucdavis.edu [michelmore@vegmail.ucdavis.edu]
 Belongs to config QH_CA.Config.192, see http://cgpdb.ucdavis.edu/
 for details.
 Plate: QH15 row: C column: 10.

FEATURES Location/Qualifiers
 source 1..284
 /organism="Helianthus annuus"
 /mol_type="mRNA"
 /cultivar="RHA601"
 /db_xref="taxon:4232"
 /clone="QH15C10"
 /lab_host="E.coli"
 /clone_lib="QH ABCDI sunflower RHA601"

/note="Vector: pBRCDNA51AB; The library was constructed
 from 11 different sources of RNA from a single genotype.
 Separate cDNAs were generated using primers that
 incorporated unique 5' and 3' tags to distinguish each
 source of RNA. cDNAs were then pooled, size-fractionated,
 directionally cloned into a custom medium-copy vector and
 transformations made with four size classes to minimize
 size bias. Details of each source of RNA and library
 construction can be obtained at http://cgpdb.ucdavis.edu/
 TAG_TISSUE=shoots environmental stress
 TAG_LIB=QH ABCDI sunflower RHA601
 TAG_SEQ=TCGCACCGCG"

ORIGIN

```

Query Match 85.6%; Score 15.4; DB 5; Length 284;  

Best Local Similarity 94.1%; Pred. No. 5.3e+03;  

Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

Qy 2 AGAACACAGATTGG 18
 |||||
 Db 82 AGAACACAGATTGG 98

RESULT 26
 BU016309 285 bp mRNA linear EST 23-AUG-2002
 LOCUS CHB12L17.yg.ab1 QH ERGHJ sunflower RHA280 Helianthus annuus cDNA
 DEFINITION clone CHB12L17, mRNA sequence.
 ACCESSION BU016309
 VERSION BU016309.1 GI:22451829
 KEYWORDS EST.
 SOURCE Helianthus annuus (common sunflower)
 ORGANISM Helianthus annuus

REFERENCE 1 (bases 1 to 285)
 Kozik, A., Michelmore, R.W., Krapp, S., Matvienko, M., Rieseberg, L.,
 Lin, H., van Damme, M., Lavelle, D., Chevaller, P., Ziegler, J.,
 Ellison, P., Kolman, J., Slabaugh, M.S., Livingston, K., Zhou, Y.,
 Lai, Z., Church, S., Jackson, L. and Bradford, K.
 Lettuce and Sunflower ESTs from the Composite Genome Project
 http://comgenomics.ucdavis.edu/
 Unpublished (2002)

COMMENT

Contact: Alexander Kozik [R.W.Michelmore]
Department of Vegetable Crops, R.W.Michelmore Lab
University of California at Davis (UCD)
Amundson Hall, UCD, Davis, CA 95616, USA
Tel: 1-(530)-742-1742
Fax: 1-(530)-752-9659
Email: akozik@ucdavis.org [michelmoreveg@mail.ucdavis.edu]
belongs to contig OH_CA_Contig192, see <http://cgpdb.ucdavis.edu/>
for details.
Plate: QHE12 row: L column: 17.
Location/Qualifiers

FEATURES

source

1..285
/organism="Helianthus annuus"
/mol_type="mRNA"
/cultivar="RHA280"
/db_xref="taxon:4232"
/clone="QHE12L17"
/lab_host="E.coli"
/clone_id="QH_EFGHJ sunflower RHA280"
/note="Vector: pBRCDNA5f1AB; The library was constructed from 11 different sources of RNA from a single genotype. Separate cDNAs were generated using primers that incorporated unique 5' and 3' tags to distinguish each source of RNA. cDNAs were then pooled, size-fractionated, directionally cloned into a custom medium-copy vector and transformations made with four size classes to minimize size bias. Details of each source of RNA and library construction can be obtained at <http://cgpdb.ucdavis.edu/>
TAG_LIB=QH_EFGHJ sunflower RHA280
TAG_SEQ=TCGCACGCGG"

ORIGIN

Query Match 85.6%; Score 15.4; DB 5; Length 285;
Best Local Similarity 94.1%; Pred. No. 5.3e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 AGAACACACAGATTCGC 18
|||||
82 AGAAAAACAGATTCGC 98

Db

RESULT 27
LOCUS BU021781 286 bp mRNA linear EST 23-AUG-2002
DEFINITION QH4e07.Y9.ab1 QH_EFGHJ sunflower RHA280 Helianthus annuus cDNA
clone QH4e07, mRNA sequence.

ACCESSION BU021781
VERSION BU021781.1 GI:22457301
KEYWORDS EST.
SOURCE Helianthus annuus (common sunflower)
ORGANISM Helianthus annuus
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; campanulids; Asterales; Asteraceae; Asteroideae;
Heliantheae; Helianthus.

REFERENCE 1 (bases 1 to 286)
Kozik, A., Michelmore, R.W., Knapp, S., Matvienko, M., Rieseberg, L.,
Litt, H., van Damme, M., Lavelle, D., Chevalier, P., Ziegler, J.,
Ellison, P., Kolman, J., Slabaugh, M.S., Livingston, K., Zhou, Y.,
Lai, Z., Church, S., Jackson, L. and Bradford, K.
Lecture and Sunflower ESTs from the Composite Genome Project
<http://comgenomics.ucdavis.edu/>
Unpublished (2002)

TITLE

JOURNAL
COMMENT Contact: Alexander Kozik [R.W.Michelmore]
Department of Vegetable Crops, R.W.Michelmore Lab
University of California at Davis (UCD)
Amundson Hall, UCD, Davis, CA 95616, USA
Tel: 1-(530)-742-1742
Fax: 1-(530)-752-9659
Email: akozik@ucdavis.org [michelmoreveg@mail.ucdavis.edu]
belongs to contig QH_CA_Contig192, see <http://cgpdb.ucdavis.edu/>
for details.

Plate: QHE4 row: e column: 07.
Location/Qualifiers

FEATURES

source

1..286
/organism="Helianthus annuus"
/mol_type="mRNA"
/cultivar="RHA280"
/db_xref="taxon:4232"
/clone="QHE4e07"
/lab_host="E.coli"
/clone_id="QH_EFGHJ sunflower RHA280"
/note="Vector: pBRCDNA5f1AB; The library was constructed from 11 different sources of RNA from a single genotype. Separate cDNAs were generated using primers that incorporated unique 5' and 3' tags to distinguish each source of RNA. cDNAs were then pooled, size-fractionated, directionally cloned into a custom medium-copy vector and transformations made with four size classes to minimize size bias. Details of each source of RNA and library construction can be obtained at <http://cgpdb.ucdavis.edu/>
TAG_LIB=QH_EFGHJ sunflower RHA280
TAG_SEQ=TCGCACGCGG"

ORIGIN

Query Match 85.6%; Score 15.4; DB 5; Length 286;
Best Local Similarity 94.1%; Pred. No. 5.3e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 AGAACACACAGATTCGC 18
|||||
82 AGAAAAACAGATTCGC 98

Db

RESULT 28
LOCUS CA485884 300 bp mRNA linear EST 14-NOV-2002
DEFINITION WHE4324_C12_F2425 wheat meiotic anther cDNA library Triticum
aestivum cDNA clone WHE4324_C12_F24, mRNA sequence.

ACCESSION CA485884
VERSION CA485884.1 GI:24979889
KEYWORDS EST.
SOURCE Triticum aestivum (bread wheat)
ORGANISM Triticum aestivum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Pooidae; Triticeae; Triticum.

REFERENCE 1 (bases 1 to 300)
Anderson, O.D., Chao, S., Crossman, C., Langridge, P., Iazo, G.R.,
Pham, J., Kausch, C.J., Sutton, T., Woo, J. and Wilson, C.
The structure and function of the expressed portion of the wheat
genomes - Meiotic anther cDNA library
Unpublished (2002)

JOURNAL
COMMENT Contact: Olin Anderson
US Department of Agriculture, Agriculture Research Service, Pacific
West Area, Western Regional Research Center
800 Buchanan Street, Albany, CA 94710, USA
Tel: 510595773
Fax: 510595818
Email: oanderson@pw.usda.gov
Sequences have been trimmed to remove vector sequence and low
quality sequence with phred score less than 20
Seq primer: SK primer
Location/Qualifiers

FEATURES

source

1..300
/organism="Triticum aestivum"
/mol_type="mRNA"
/cultivar="Chinese Spring"
/db_xref="taxon:4565"
/clone="WHE4324_C12_F24"
/issue_type="Anther"
/dev_stage="Meiotic stages pre-meiosis-metaphase I"
/lab_host="E. coli DH10B"
/clone_lib="wheat meiotic anther cDNA library"

/note="Vector: pSPORT1, Site 1: SalI; Site 2: NotI; Plants were grown in a glasshouse. Another meiotic stage was determined by removing anthers from individual primary florets. One anther was sacrificed for microscopic staging, and if determined to be between (and including) meiotic stages pre-meiosis and metaphase I, the remaining two anthers were collected and pooled for library construction. The tissue, total RNA, and poly(A) RNA were prepared. cDNA synthesised, and directionally ligated into pSPORT1 by Tim Sutton in the P Langridge Lab at the Department of Plant Science, University of Adelaide, Waite Campus, Australia. Average insert size 1.5Kb. Plasmid DNA preparations and DNA sequencing were performed in the OD Anderson lab (all other authors)."

ORIGIN

Query Match 85.6%; Score 15.4; DB 6; Length 300;
Best Local Similarity 94.1%; Pred. No. 5.3e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 AGAACACACAGATTGCG 18
|||||
Db 10 AGAACACACAGATTGCG 26

RESULT 29
AM864697/c 304 bp mRNA linear EST 22-MAY-2000
LOCUS PM2-SN0018-160300-001-d01 SN0018 Homo sapiens cDNA, mRNA sequence.
DEFINITION
ACCESSION AM864697
VERSION AM864697.1 GI:7998747
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 304)
Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R., Nagai, M.A., da Silva, M. U.T., Zago, N.A., Boridin, S., Costa, F.F., Goldman, G.H., Carvalho, A.P., Matsukuma, A., Bata, G.S., Simpson, D.H., Brunstein, A., deoliveira, P.S., Bucher, P., Jorgensen, C.V., O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and Simpson, A.J.

Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
10737800
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil

COMMENT

Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the PAPESP/LICP Human Cancer Genome Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=kt2=PM2-SN0018-160300-001-d01&t3=2000-03-16&t4=1)
Seq primer: puc 18 forward
High quality sequence stop: 304.

FEATURES

Location/Qualifiers
1 304

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="SN0018"
/note="Organ: stomach normal; Vector: puc18; Site 1: SmaI; Site 2: SmaI; A mini-library was made by cloning products derived from ORFESTS PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research)

profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

ORIGIN

Query Match 85.6%; Score 15.4; DB 2; Length 304;
Best Local Similarity 94.1%; Pred. No. 5.3e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 AGAACACACAGATTGCG 18
|||||
Db 63 AGAACACACAGATTGCG 47

RESULT 30
AV216565
LOCUS AV216565 312 bp mRNA linear EST 30-OCT-1999
DEFINITION AV216565 RIKEN full-length enriched, ES cells Mus musculus cDNA clone 2410164G16 3' similar to X57708 M. musculus RNA for pH 34, mRNA sequence.
ACCESSION AV216565
VERSION AV216565.1 GI:6157410
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus

REFERENCE
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 312)
Fukuda, S., Fukunishi, Y., Hara, A., Hayatsu, N., Hirozane, T., Hori, F., Ishii, Y., Ishikawa, T., Iton, M., Izawa, M., Kadota, K., Kagawa, I., Kai, C., Kawai, J., Kikuchi, N., Kojima, Y., Koya, S., Kusakabe, M., Matsuyama, T., Miki, R., Mizuno, Y., Nakamura, M., Oda, H., Okazaki, Y., Owa, C., Ozawa, Y., Saito, H., Sano, M., Sato, K., Shibata, K., Shibata, Y., Shigemoto, Y., Shiraki, T., Sogabe, Y., Suganara, Y., Suzuki, H., Suzuki, H., Takahashi, F., Tateno, M., Tomioka, N., Tsunoda, Y., Watabiki, A., Watanabe, S., Yamamura, T., Yasunishi, A., Yokota, T., Yoshiki, A., Yoshino, M., Yamatsu, M., and Hayashizaki, Y.
RIKEN Mouse ESTs (Kono, H., et al. 1999)
Unpublished (1999)

COMMENT

Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suenho-cho, Tsunumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-resgsc.riken.jp, URL: http://genome.gsc.riken.jp/
Sasaki, N., Izawa, M., Watabiki, M., Ozawa, K., Tanaka, T., Yoneda, Y., Matsuyama, S., Carninci, P., Muramatsu, M., Okazaki, Y. and Hayashizaki, Y.
Transcriptional sequencing: A method for DNA sequencing using RNA polymerase. Proc. Natl. Acad. Sci. U.S.A. 95 (7), 3455-3460 (1998)
Itoh, M., Kiteunai, T., Akiyama, J., Shibata, K., Izawa, M., Kawai, J., Tomaru, Y., Carninci, P., Shibata, Y., Ozawa, Y., Muramatsu, M., Okazaki, Y. and Hayashizaki, Y.
Automated filtration-based high-throughput plasmid preparation system. Genome Res. 9 (5), 463-470 (1999)
Carninci, P. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning. Methods Enzymol. 303, 19-44 (1999)
Please visit our web site (http://genome.rtc.riken.go.jp) for further details.

FEATURES

Location/Qualifiers
1 312

/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone_lib="2410164G16"
/cell_type="ES cells"
/lab_host="SOLAR"
/clone_lib="RIKEN full-length enriched, ES cells"

/note="Site 1: XhoI; Site 2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genetic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5', GAGAGAGAGATTCGATTCGATTAATTAATATCCGCCGCCGCC 3'] cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rct = 5.0 and subtraction to Rct = 25.0. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGAGAGATTCGATTCGATTAATTAATATCCGCCGCCGCC 3']."

ORIGIN

Query Match 85.6%; Score 15.4; DB 1; Length 312;
Best Local Similarity 94.1%; Pred. No. 5.3e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 AGAACAACAAGATTGCG 18
DB 28 AAAACAACAAGATTGCG 44

RESULT 31

AA787371 340 bp mRNA linear EST 31-JUL-1998
LOCUS n2g08a1.t1 Aspergillus nidulans 24hr asexual developmental and
DEFINITION vegetative cDNA lambda zap library Emericella nidulans cDNA clone
n2g08a1, mRNA sequence.

ACCESSION AA787371.1 GI:2847602
VERSION AA787371.1
KEYWORDS Emericella nidulans (anamorph: Aspergillus nidulans)
SOURCE Emericella nidulans
ORGANISM Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
Eurotiales; Trichocomaceae; Emericella.

REFERENCE 1 (bases 1 to 340)
AUTHORS Kudler, D., Gray, J., Hausner, J., Lai, H., Martin, W., Aramayo, R.,
Prade, R. and Roe, B.
TITLE An Aspergillus nidulans EST Database
JOURNAL Unpublished (1998)
COMMENT Contact: Bruce A. Roe, University of Oklahoma, broeou.edu
Department of Chemistry and Biochemistry
Advanced Center for Genome Technology, University of Oklahoma
620 Parrington Oval, Norman, OK 73019, USA
Tel: 405 325 4912
Fax: 405 325 7762
Email: broeou.edu
We anticipate the future release of the cDNA clones to the Fungal
Genetics Stock Center
High quality sequence stop: 326.

FEATURES

source

1..340
Location/Qualifiers
/organism="Emericella nidulans"
/mol_type="mRNA"
/strain="FGSC A26"
/db_xref="taxon:162425"
/clone="n2g08a1"
/tissue_type="vegetative mycelia, asexual structures"
/clone_lib="Aspergillus nidulans 24hr asexual
developmental and vegetative cDNA lambda zap library"
/note="Vector: pBluescript SK-; Site 1: EcoRI; Site 2:
XhoI; 5' end of cDNA cloned into EcoRI site of pBluescript
3' end of cDNA cloned into XhoI site of pBluescript"

ORIGIN

Query Match 85.6%; Score 15.4; DB 1; Length 340;
Best Local Similarity 94.1%; Pred. No. 5.4e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GAGAACAACAAGATTGCG 17

Db 278 GAGAACAACAAGATTGCG 294

RESULT 32
LOCUS AW447131 351 bp mRNA linear EST 25-APR-2001
DEFINITION 88282 MARC 1BOV Bos taurus cDNA 5', mRNA sequence.
ACCESSION AW447131
VERSION AW447131.1 GI:6988918
KEYWORDS EST.
SOURCE Bos taurus (cow)
ORGANISM Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovinae; Bos.

REFERENCE 1 (bases 1 to 351)
AUTHORS Smith, T.P.L., Gross, W.M., Freking, B.A., Roberts, A.J., Stone, R.T.,
Caas, E., Wray, J.E., White, J., Cho, J., Fahrenkrug, S.C.,
Bennett, G.L., Heaton, M.P., Jaeger, W.W., Rohrer, G.A.,
Chitko-McKown, C.G., Petree, G., Holt, I., Karaycheva, S., Liang, F.,
Quackenbush, J., and Keele, J.W.
TITLE Sequence evaluation of four pooled-tissue normalized bovine cDNA
libraries and construction of a gene index for cattle
JOURNAL Genome Res. 11 (4), 626-630 (2001)
MEDLINE 21180013
PUBMED 11282978

COMMENT Contact: Smith TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smith@email.marc.usda.gov
Single pass sequencing. Bases called and trimmed with phred
v0.980904.e. Vector identified by cross_match with the -minscore 20
and -mismatch 12 options.

PCR PRIMERs
FORWARD: AGGAACAAGTATGACCAT
BACKWARD: GTTTCAGTCAGCAGC
Plate: 63 row: 1 column: 5
Seq primer: ATTTCGTGACACTATG.
Location/Qualifiers
1..351
/organism="Bos taurus"
/mol_type="mRNA"
/db_xref="taxon:9913"
/tissue_type="pooled"
/lab_host="PDH10B"
/clone_lib="MARC 1BOV"
/note="Vector: pCMV SPORT6; Site 1: NotI; Site 2: SalI;
Library made from pooled tissue from lymph node, ovary,
fat, hypothalamus, and pituitary."

FEATURES

source

1..351
Location/Qualifiers
/organism="Bos taurus"
/mol_type="mRNA"
/db_xref="taxon:9913"
/tissue_type="pooled"
/lab_host="PDH10B"
/clone_lib="MARC 1BOV"
/note="Vector: pCMV SPORT6; Site 1: NotI; Site 2: SalI;
Library made from pooled tissue from lymph node, ovary,
fat, hypothalamus, and pituitary."

ORIGIN

Query Match 85.6%; Score 15.4; DB 2; Length 351;
Best Local Similarity 94.1%; Pred. No. 5.4e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GAGAACAACAAGATTGCG 17
DB 74 GAGAACAACAAGATTGCG 90

RESULT 33
LOCUS AU257195 365 bp mRNA linear EST 25-APR-2002
DEFINITION AU257195 3'-directed mouse cDNA library Mus musculus cDNA clone
BE00009976 3', mRNA sequence.

ACCESSION AU257195
VERSION AU257195.1 GI:20321577
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 365)
AUTHORS Kato, K. and Matoba, R.
TITLE Generation of expressed sequence tags from mouse brain
JOURNAL Unpublished (2002)
COMMENT Contact: Kikuya Kato
Graduate School of Biological Sciences
Nara Institute of Science and Technology
8916-5 Takayama, Ikoma, Nara 630-0101, Japan
Tel: 81-743-72-5581
Fax: 81-743-72-5589
Email: kkatob@is.nara.ac.jp,
URL: http://love2.aist-nara.ac.jp/BEI/index.html.

FEATURES
source
1..365
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="BBD0009976"
/tissue_type="brain"
/clone_lib="3-directed mouse cDNA library"

ORIGIN
Query Match 85.6%; Score 15.4; DB 1; Length 365;
Best Local Similarity 94.1%; Pred. No. 5.4e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

CY 2 AGACACACAGATTGCG 18
|||||
Db 306 AGACACACAAATTCGC 322
|||||

RESULT 34
LOCUS B2651766 366 bp DNA linear GSS 29-JAN-2003
DEFINITION OGANP82C ZM_0.7_1.5_KB zea mays genomic clone ZMWBMA0102M20,
genomic survey sequence.
ACCESSION B2651766
VERSION B2651766.1 GI:28118555
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 366)
Whiteaw, C.A., Quackenbush, J., Van Aken, S., Uterback, T.,
Resnick, A., Fraser, C.M., Budiman, M.A., Bedell, J.A., Rohlfing, T.,
Citek, R.W., Nuneberg, A., Robbins, D. and Lakey, N.
Consortium for Maize Genomics
Unpublished (2002)
Other GSSs: OGANP82TM
Contact: Cathy Whiteaw
TIGR Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whiteaw@tigr.org
Seq primer: TF
Class: sheared ends.

FEATURES
source
1..366
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMWBMA0102M20"
/clone_lib="ZM_0.7_1.5_KB"
/note="Vector: pGCSK-; Site 1: HincII; 0.7-1.5 kb
methylation filtered cDNA library"

ORIGIN

Query Match 85.6%; Score 15.4; DB 8; Length 366;
Best Local Similarity 94.1%; Pred. No. 5.4e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

CY 2 AGACACACAGATTGCG 18
|||||
Db 107 AGACACACAAATTCGC 91
|||||

RESULT 35
LOCUS BF851239 367 bp mRNA linear EST 16-JAN-2001
DEFINITION IL5-EN0086-281100-291-g03 EN0086 Homo sapiens cDNA, mRNA sequence.
ACCESSION BF851239
VERSION BF851239.1 GI:12238401
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 367)
Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,
Nagai, M.A., da Silva, M. Jr., Zago, M.A., Bordin, S., Costa, F.F.,
Goldman, G.H., Carvalho, A.F., Matsukuma, A., Bala, G.S., Simpson, D.H.,
Brunstein, A., de Oliveira, P.S., Bucher, P., Jongeneel, C.V.,
O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and
Simpson, A.J.
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
20202663
10737800
PUBMED
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the PAPER/P/PCR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?vl=IL5&2=IL5-EN0086-
281100-291-g03&t3=2000-11-28&t4=1)
Seq primer: puc 18 forward
High quality sequence stop: 336.
Location/Qualifiers
1..367
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="EN0086"
/note="Organ: lung normal; Vector: puc18; Site 1: SmaI;
Site 2: SmaI; A mini-library was made by cloning products
derived from ORSITES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."

ORIGIN
Query Match 85.6%; Score 15.4; DB 2; Length 367;
Best Local Similarity 94.1%; Pred. No. 5.4e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

CY 2 AGACACACAGATTGCG 18
|||||
Db 218 AGACACACAAATTCGC 234
|||||

RESULT 36
AY440562/c

TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cga@bbs-rcmail.nih.gov
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 877 Std Error: 0.00
Seq primer: -40m13 fwd. ET from Amersham
High quality sequence stop: 389.
Location/Qualifiers
1..400
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:1625908"
/lab_host="DH10B"
/clone_lib="Soares_NFL_T_GBC_S1"
/note="Organ: pooled; Vector: PT7T3D-Pac (Pharmacia) with
a modified polylinker; Site 1: Not I; Site 2: Eco RI;
Equal amounts of plasmid DNA from three normalized
libraries (fetal lung NDH19W, testis NBT, and B-cell
NCI CGAP GCBI) were mixed, and ss circles were made in
vitro. Following HAP purification, this DNA was used as
tracer in a subtractive hybridization reaction. The driver
was PCR-amplified cDNAs from pools of 5,000 clones made
from the same 3 libraries. The pools consisted of
I.M.A.G.E. clones 257480-332087, 682632-687239,
726408-726711, and 729096-731399. Subtraction by Bento
Soares and M. Fatima Ronaldo."

ORIGIN
Query Match 85.6%; Score 15.4; DB 1; Length 400;
Best Local Similarity 94.1%; Pred. No. 5.4e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 AGAACAACAGATTCG 18
|||
247 AGCACAACAGATTCG 263

Db

RESULT 39
CG722358/c 402 bp DNA linear GSS 20-OCT-2003
LOCUS 119071F06.1BL Y1 1119 - Rescuedu Grid AA Zee may's genomic, genomic
DEFINITION survey sequence.
ACCESSION CG722358
VERSION CG722358.1 GI:37757127
KEYWORDS GSS.
SOURCE Zee may's
ORGANISM Zee may's
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoidae; Andropogoneae; Zee.
1 (bases 1 to 402)
Walbot, V.
Maize genomic sequences found using engineered Rescuedu transposon
Unpublished (2001)
Contact: Walbot V
Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 725 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Very probable ligation site of ends cut by single endonuclease.
Reverse complemented post-ligation sequence from source sequence.
Plate: 119071 row: 19
Class: transposon-tagged.
Location/Qualifiers
1..402
/organism="Zee may's"
/mol_type="genomic DNA"

FEATURES
source

/cultivar="mixed background W23/Al86/B73/K55"
/db_xref="taxon:4577"
/tissue_type="leaf"
/dev stage="adult"
/lab_host="DH10B"
/clone_lib="1119 - Rescuedu Grid AA"
/note="Organ: leaf; Vector: Rescuedu (engineered from
pBluescript backbone); Site 1: BamHI, Site 2: BglII;
Rescuedu is a 4.9 kb, modified maize Mu transposon
designed to allow plasmid rescue from total genomic DNA.
Mu elements insert preferentially into transcription
units. For more information on Rescuedu, go to the web
site 'www.zmnd.iastate.edu' and follow the links for
'Rescuedu'. Grid AA was grown at UC San Diego in 2002. DNA
was extracted from leaf strips, double digested using
BamHI and BglII, and ligated to form circular plasmids.
DH10B cells were transformed and then screened on LB
plates with ampicillin."

ORIGIN
Query Match 85.6%; Score 15.4; DB 9; Length 402;
Best Local Similarity 94.1%; Pred. No. 5.4e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GAGAACAACAGATTCG 17
|||
223 GAGACAACAGATTCG 213

Db

RESULT 40
AI280419 406 bp mRNA linear EST 29-JAN-1999
LOCUS g195d08.x1 Soares_NHMPu S1 Homo sapiens cDNA clone IMAGE:1880079
DEFINITION 3' similar to TR.F70193 F70193 INTEGRAL MEMBRANE GLYCOPROTEIN. ;,
mRNA sequence.
ACCESSION AI280419
VERSION AI280419.1 GI:3918652
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
1 (bases 1 to 406)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cga@bbs-rcmail.nih.gov
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 621 Std Error: 0.00
Seq primer: -40UP from Gibco
High quality sequence stop: 406.
Location/Qualifiers
1..406
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:1880079"
/tissue_type="Pooled human melanocyte, fetal heart, and
pregnant uterus"
/lab_host="DH10B"
/clone_lib="Soares_NHMPu_S1"
/note="Organ: mixed (see below); Vector: PT7T3D-Pac
(Pharmacia) with a modified polylinker; Site 1: Not I;
Site 2: Eco RI; Equal amounts of plasmid DNA from three
normalized libraries (melanocyte 2NBH, pregnant uterus
NBHPU, and fetal heart NDH19W) were mixed, and ss circles
were made in vitro. Following HAP purification, this DNA
was used as tracer in a subtractive hybridization
reaction. The driver was PCR-amplified cDNAs from pools of
5,000 clones made from the same 3 libraries. The pools

ORIGIN
Query Match 85.6%; Score 15.4; DB 1; Length 406;
Best Local Similarity 94.1%; Pred. No. 5.5e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 AGAACACACAGATTCCG 18
DB 245 AGCACACACAGATTCCG 261

RESULT 41
LOCUS A1220277 409 bp mRNA linear EST 30-NOV-1998
DEFINITION g973d03.x1 Soares NFL T GBC S1 Homo sapiens CDNA clone
IMAGE:1840805 3' similar to TR:P70193 P70193 INTEGRAL MEMBRANE
GLYCOPROTEIN.; mRNA sequence.

ACCESSION A1220277
VERSION A1220277.1 GI:3802480
KEYWORDS EST.
SOURCE Homo sapiens (human)

ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 409)
AUTHORS NCI-CCAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.lnl.gov) for further information.
Insert Length: 1154 Std Error: 0.00
Seq primer: -40UP from Gibco
High quality sequence stop: 399.

FEATURES
source
1..409
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:1840805"
/lab_host="DH10B"
/clone_lib="Soares NFL T GBC S1"
/note="Organ: pooled; Vector: pT73D-Pac (Pharmacia) with
a modified polylinker; Site 1: Not I; Site 2: Eco RI;
Equal amounts of plasmid DNA from three normalized
libraries (fetal lung NDHL19W, testis NT, and B-cell
NCI CGAP GCB1) were mixed, and ss circles were made in
vitro. Following HAP purification, this DNA was used as
tracer in a subtractive hybridization reaction. The driver
was PCR-amplified cDNAs from pools of 5,000 clones made
from the same 3 libraries. The pools consisted of
I.M.A.G.E. clones 297480-302087, 682632-687239,
726408-728711, and 729096-731399. Subtraction by Bento
Soares and M. Fatima Bonaldo."

ORIGIN

Query Match 85.6%; Score 15.4; DB 1; Length 409;
Best Local Similarity 94.1%; Pred. No. 5.5e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 AGAACACACAGATTCCG 18
DB 170 AGCACACACAGATTCCG 166

RESULT 42
LOCUS AA907628 413 bp mRNA linear EST 10-JUN-1998
DEFINITION cm10f08.s1 Soares NFL T GBC S1 Homo sapiens CDNA clone

IMAGE:1540647 3' similar to TR:P70193 P70193 INTEGRAL MEMBRANE
GLYCOPROTEIN.; mRNA sequence.
AA907628
VERSION A1220277.1 GI:3043088
KEYWORDS EST.
SOURCE Homo sapiens (human)

ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 413)
AUTHORS NCI-CCAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.lnl.gov) for further information.
Insert Length: 606 Std Error: 0.00
Seq primer: -40m13 fwd. ET from Amersham
High quality sequence stop: 274.

FEATURES
source
1..413
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:1540647"
/lab_host="DH10B"
/clone_lib="Soares NFL T GBC S1"
/note="Organ: pooled; Vector: pT73D-Pac (Pharmacia) with
a modified polylinker; Site 1: Not I; Site 2: Eco RI;
Equal amounts of plasmid DNA from three normalized
libraries (fetal lung NDHL19W, testis NT, and B-cell
NCI CGAP GCB1) were mixed, and ss circles were made in
vitro. Following HAP purification, this DNA was used as
tracer in a subtractive hybridization reaction. The driver
was PCR-amplified cDNAs from pools of 5,000 clones made
from the same 3 libraries. The pools consisted of
I.M.A.G.E. clones 297480-302087, 682632-687239,
726408-728711, and 729096-731399. Subtraction by Bento
Soares and M. Fatima Bonaldo."

ORIGIN

Query Match 85.6%; Score 15.4; DB 1; Length 413;
Best Local Similarity 94.1%; Pred. No. 5.5e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 AGAACACACAGATTCCG 18
DB 244 AGCACACACAGATTCCG 260

RESULT 43
LOCUS BM173099 419 bp mRNA linear EST 04-DEC-2001
DEFINITION 900453 Avicennia marina leaf CDNA library Avicennia marina CDNA
clone Am900453 5' similar to unknown protein (AB025633) of
Arabidopsis thaliana, mRNA sequence.

ACCESSION BM173099
VERSION BM173099.1 GI:17312662
KEYWORDS EST.
SOURCE Avicennia marina
ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; Lamiales; Acanthaceae; Acanthaceae incertae
sedis; Avicennia.

REFERENCE 1 (bases 1 to 419)
AUTHORS Parani M., Mehta P., Sivaprakash K.R. and Parida A.
JOURNAL Unpublished (2000)
COMMENT Contact: Parani M / Parida A
Department of Plant Molecular Biology

M. S. Swaminathan Research Foundation
 III Cross Street, Taramani Institutional Area, Chennai 600 113,
 India
 Tel: 91-44-2351319
 Fax: 91-44-2351319
 Email: mangrovegenes@msrf.res.in
 BLAST search in BLASTX (Non-redundant) using default parameters as
 on November 6th 2001
 Seq primer: M13 Reverse.
 Location/Qualifiers

FEATURES
 source
 1..419

/organism="Avicennia marina"
 /mol_type="mRNA"
 /strain="Pichavaram"
 /db_xref="taxon:82927"
 /clone="Am900453"
 /sex="Hermaphrodite"
 /tissue_type="leaf"
 /dev_stage="1 year old plant"
 /clone_lib="Avicennia marina leaf cDNA library"
 /note="A one-year-old plant from the natural mangrove
 habitat in Pichavaram, Tamil Nadu, India was collected
 and brought to the M.S. Swaminathan Research Foundation.
 The whole plant was treated with the nutrient solution
 supplemented with 500 mM NaCl for 48 hours. Poly(A⁺) mRNA
 from the leaf tissue was purified and cDNA was prepared
 using Superscript Lambda System (Life Technologies, Cat.
 No. 19643-014). The cDNAs were size fractionated over
 SizeSep-400 spun column (Amersham-Pharmacia Biotech Cat.
 No. 27-5105-01) and cloned in 5'Sal I - 3' Not I of
 pSPORT1 (Life Technologies Cat. No. 15383-011). The ESTs
 were sequenced from the 5' end using M13/pUC18 reverse
 primer in an Automated Sequencer (ABI310, Applied
 Biosystems) and submitted after editing to remove the
 vector and adapter sequences. BLAST search in BLASTX
 (Non-redundant) was carried out on 7th August, 2000 using
 default parameters and the results reported under
 putative identification/comment"

ORIGIN

Query Match 85.6%; Score 15.4; DB 4; Length 419;
 Best Local Similarity 94.1%; Pred. No. 5.5e+03;
 Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GAGACACACAGATTTCG 17
 |||
 82 GAGGACACACAGATTTCG 98

RESULT 44
 AA812092 428 bp mRNA linear EST 19-FEB-1998
 LOCUS AA812092
 DEFINITION cb040a06.61 NCI CGAP GCB1 Homo sapiens cDNA clone IMAGE:133810 3'
 similar to TR:P70193 P70193 MEMBRANE GLYCOPROTEIN.; mRNA
 sequence.

ACCESSION AA812092
 VERSION AA812092.1 GI:2881703
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 428)
 NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 Unpublished (1997)
 CONTACT: Robert Strausberg, Ph.D.
 Email: cgapdb-remail.nih.gov
 Tissue Procurement: Louis M. Staudt, M.D., Ph.D., David Allman,
 Ph.D., Gerald Marti, M.D.
 cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
 Bernaldo, Ph.D.

JOURNAL
 COMMENT

cDNA Library Arrayed by: Greg Lennon, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bdrp/image/image.html
 Insert length: 1302 Std Error: 0.00
 Seq primer: -40m3 fwd. ET from Amersham.
 Location/Qualifiers

FEATURES
 source
 1..428

/organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:133810"
 /tissue_type="germinal center B cell"
 /lab_host="DH10B"
 /clone_lib="NCI CGAP GCB1"
 /note="Vector: pT7T3D-Pac (pharmacia) with a modified
 polylinker. Site 1: Not I; Site 2: Eco RI; 1st strand cDNA
 was prepared from human consillar cells enriched for
 germinal center B cells by flow sorting (CD20⁺, IgD⁻),
 provided by Dr. Louis M. Staudt (NCI), Dr. David Allman
 (NCI) and Dr. Gerald Marti (CBER). cDNA synthesis was
 primed with a Not I - oligo(dT) primer
 [5'-TGTTACCAATCTGAAGTGGAGCGCGCCCTCATTTTCTTTTCTTTT-3'
]. Double-stranded cDNA was ligated to Eco RI adaptors
 (pharmacia), digested with Not I and cloned into the Not I
 and Eco RI sites of the modified pT7T3 vector. Library
 went through one round of normalization, and was
 constructed by Bento Soares and M. Fatima Bernaldo."

ORIGIN

Query Match 85.6%; Score 15.4; DB 1; Length 428;
 Best Local Similarity 94.1%; Pred. No. 5.5e+03;
 Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 AGACACACAGATTTCG 18
 |||
 Db 213 AGCACACACAGATTTCG 229

RESULT 45
 BZ769778 430 bp DNA linear GSS 13-MAR-2003
 LOCUS BZ769778/c
 DEFINITION SALK_142711.47.00.x Arabidopsis thaliana TDNA insertion lines
 Arabidopsis thaliana genomic clone SALK_142711.47.00.x, genomic
 survey sequence.

ACCESSION BZ769778
 VERSION BZ769778.1 GI:28943462
 KEYWORDS GSS.
 SOURCE Arabidopsis thaliana (thale cress)
 ORGANISM Arabidopsis thaliana

REFERENCE 1 (bases 1 to 430)
 Alonso, J.M., Leisse, T.J., Barajas, P., Chen, H., Cheuk, R.,
 Gadrinab, C., Jeske, A., Karnes, M., Kim, C.J., Parker, H., Prednis, L.,
 Shinn, P., Zimmerman, J. and Ecker, J.R.
 A Sequence-Indexed Library of Insertion Mutations in the
 Arabidopsis Genome
 Unpublished (2001)
 CONTACT: Joseph R. Ecker
 Salk Institute Genomic Analysis Laboratory (SIGAL)
 The Salk Institute for Biological Studies
 10010 N. Torrey Pines Road, La Jolla, CA 92037, USA
 Tel: 858 453 4100 x1752
 Fax: 858 558 6379
 Email: ecker@salk.edu

JOURNAL
 COMMENT
 This is single pass sequence recovered from the left border of
 TDNA.
 Class: TDNA tagged.
 Location/Qualifiers

FEATURES
 source
 1..430

/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/ecotype="Col-0"
/db_xref="taxon:3702"
/clone_lib="Arabidopsis thaliana TDNA insertion lines"
/note="PCR was performed on Arabidopsis thaliana lines each of which contains one or more TDNA insertion elements. The resultant fragment for each line was directly sequenced to determine the genomic sequence at the site of insertion. Details of the protocols used can be found at http://signal.salk.edu/tdna_protocols.html"

ORIGIN

Query Match 85.6%; Score 15.4; DB 8; Length 430;
Best Local Similarity 94.1%; Pred. No. 5.5e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GAGAACACAGATTTCG 17
|||||
Db 284 GAGAACACAGATTTCG 268

RESULT 46
BP645308 431 bp mRNA linear EST 27-JUN-2004
LOCUS BP645308 RAF1.9 Arabidopsis thaliana cDNA clone RAF1.9-68-F13 3',
DEFINITION mRNA sequence.

ACCESSION BP645308
KEYWORDS BP645308.1 GI:49296778
SOURCE EST.
ORGANISM Arabidopsis thaliana (thale cress)

Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
1 (bases 1 to 431)

REFERENCE
AUTHORS Seki, M., Narusaka, M., Kamiya, A., Ishida, J., Satou, M., Sakurai, T.,
Nakaizuma, M., Enju, A., Akiyama, K., Oono, Y., Muramatsu, M.,
Hayashizaki, Y., Kawai, J., Carninci, P., Itoh, M., Ishii, Y.,
Aizawa, T., Shibata, K., Shingawa, A. and Shinozaki, K.,
Functional annotation of a full-length Arabidopsis cDNA collection
Science 296 (5565), 141-145 (2002)

COMMENT

11910074
Contact: Motoaki Seki
Plant Functional Genomics Research Group
RIKEN Genomic Sciences Center
3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan
Tel: 81-298-36-4359
Fax: 81-298-36-9060
Email: mseki@tc.riken.go.jp
reversed clone; please visit our web site
(<http://pfigweb.gsc.riken.go.jp/>) for further details.
Location/Qualifiers

FEATURES

1..431
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/clone="RAF1.9-68-F13"
/issue_type="mixture of silique and flower"
/lab_host="DH10B"
/clone_lib="RAF1.9"
/note="Site_1: BamHI, Site_2: SalI, Subtraction library"

ORIGIN

Query Match 85.6%; Score 15.4; DB 5; Length 431;
Best Local Similarity 94.1%; Pred. No. 5.5e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 AGAACACACAGATTTCG 18
|||||
Db 92 AGAACACACAGATTTCG 108

RESULT 47
BG543327/c 433 bp mRNA linear EST 01-MAY-2002
LOCUS BG543327/c
DEFINITION E0832 Chinese cabbage etiolated seedling library Brassica rapa
subsp. pekinensis cDNA clone E0832, mRNA sequence.

ACCESSION BG543327
VERSION BG543327.1 GI:20374307
KEYWORDS EST.
SOURCE Brassica rapa subsp. pekinensis
ORGANISM Brassica rapa subsp. pekinensis
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Brassica.

REFERENCE
AUTHORS Ryu, S.-H., Kang, U.-S., Kang, C.-h., Kim, C.-Y., Choi, Y.-J., Lee, S.-H.,
Bak, J.-D., Lee, S.-Y., Cho, M.-J. and Lim, C.-O.
Expressed Sequence Tags of Chinese Cabbage Etiolated Seedling cDNA
Unpublished (2001)
Contact: Lim, C.O.
Plant Molecular Biology & Biotechnology Research Centre
Gyeongsang National University
#900 Gazwa-dong, Jinju 660-701, Korea
Tel: 82 55 751 6255
Fax: 82 55 759 9363
Email: collim@ongae.gsnu.ac.kr
Seq primer: 17.

FEATURES

Location/Qualifiers
1..433
/organism="Brassica rapa subsp. pekinensis"
/mol_type="mRNA"
/cultivar="Jangwon"
/sub_species="pekinensis"
/db_xref="taxon:51351"
/clone="E0832"
/issue_type="Etiolated seedling"
/lab_host="XL-1 Blue"
/clone_lib="Chinese cabbage etiolated seedling library"
/note="Vector: pSPORT 1; Site_1: Sal I; Site_2: Not I"

ORIGIN

Query Match 85.6%; Score 15.4; DB 4; Length 433;
Best Local Similarity 94.1%; Pred. No. 5.5e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 AGAACACACAGATTTCG 18
|||||
Db 159 AGAACACACAGATTTCG 143

RESULT 48
BM360109/c 437 bp mRNA linear EST 09-JAN-2002
LOCUS BM360109/c
DEFINITION GA_Ea0027G15+ Gossypium arboreum 7-10 dpa fiber library Gossypium
arboreum cDNA clone GA_Ea0027G15+, mRNA sequence.

ACCESSION BM360109
VERSION BM360109.1 GI:18100855
KEYWORDS EST.
SOURCE Gossypium arboreum
ORGANISM Gossypium arboreum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Malvales; Malvaceae; Malvoideae; Gossypium.

REFERENCE
AUTHORS Wing, R.A., Fritsch, D., Yu, Y., Main, D., Rambo, T., Simmons, J.,
Henry, D., Wood, T.C., Leslie, A. and Wilkins, T.A.
An integrated analysis of the genetics, development, and evolution
of the cotton fiber
Unpublished (2000)
Contact: Wing RA
Clemson University Genomics Institute
Clemson University

100 Jordan Hall, Clemson, SC 29634, USA
Tel: 864 656 7288
Fax: 864 656 4293
Email: rwing@clemson.edu
Total High Quality bases = 129
Seq primer: TAATGCACTCAGCTAGG
High quality sequence start: 15
High quality sequence stop: 395.
Location/Qualifiers

FEATURES

source

1. 437
/organism="Gossypium arboreum"
/mol_type="mRNA"
/strain="AKA"
/cultivar="8400"
/db_xref="taxon:29729"
/clone="GA_EA027G15r"
/tissue_type="Fibers isolated from bolls harvested 7-10 dpa"
/lab_host="E. coli"
/clone_lib="Gossypium arboreum 7-10 dpa fiber library"
/note="Vector: pBK-CMV; Site_1: EcoRI; Site_2: XhoI"

ORIGIN

Query Match 85.6%; Score 15.4; DB 4; Length 437;
Best Local Similarity 94.1%; Pred. No. 5.5e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 AGACACACAGATTGCG 18
|||||
DB 39 AGACACACAGATTGCG 23

RESULT 49
AA610098 438 bp mRNA linear EST 30-SEP-1997
LOCUS at19e07.s1 Soares_total_fetus_Nb2HF8_9w Homo sapiens cDNA clone
DEFINITION IMAGE:1032132 3' similar to SW:SLIT_DROME P24014 SLIT PROTEIN
PRECUSOR.1; mRNA sequence.

ACCESSION AA610098
VERSION AA610098.1 GI:2458526
KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 438)
Hillier, L., Allen, M., Bowles, J., Dubuque, T., Gaisel, G., Jost, S.,
Krizman, D., Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M.,
Martin, J., Moore, B., Scheinberg, K., Stepcoe, X., Tan, F.,
Thaising, B., White, Y., Wylie, T., Waterston, R. and Wilson, R.

AUTHORS

WashU-NCI human EST Project
Unpublished (1997)

JOURNAL

CONTACT: Wilson RK

COMMENT

Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810

Email: est@watson.wustl.edu
This clone is available royalty-free through LML; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Possible reversed clone: similarity on wrong strand
Seq primer: -40m3 fwd. ET from Amersham
High quality sequence stop: 399.
Location/Qualifiers

FEATURES

source

1. 438
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:1032132"
/dev_stage="8-9 weeks"
/lab_host="DH10B"
/clone_lib="Soares_total_fetus_Nb2HF8_9w"

/note="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
was prepared from mRNA obtained from pooled 8-9 week
(total) fetus material with a Not I - oligo(dT) primer [5'
TGTAACCAATCGAATGGAGCGCGCGCTATATTTTTTTTTTTT 3'].
Double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT73 vector. Library
went through one round of normalization, and was
constructed by Bento Soares and M. Fatima Bonaldo."

ORIGIN

Query Match 85.6%; Score 15.4; DB 1; Length 438;
Best Local Similarity 94.1%; Pred. No. 5.5e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 AGACACACAGATTGCG 18
|||||
DB 244 AGACACACAGATTGCG 260

RESULT 50

BU801478 438 bp mRNA linear EST 26-MAY-2004
LOCUS BU801478 unpublished oligo-capped cDNA library, stage I4
DEFINITION Caenorhabditis elegans cDNA clone yk144912 5', mRNA sequence.

ACCESSION BU801478
VERSION BU801478.1 GI:47705145
KEYWORDS EST.
SOURCE Caenorhabditis elegans
ORGANISM Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida;
Rhabditidae; Rhabditidae; Peloderinae; Caenorhabditis.

REFERENCE 1 (bases 1 to 438)
Kohara, Y., Shin, I.T., Thierry-Mieg, J., Thierry-Mieg, D., Suzuki, Y.
and Sugano, S.
A complementary view of the C. elegans genome

AUTHORS

Unpublished (2002)
Contact: Tadao Shin-I
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855

Email: tshin@genes.nig.ac.jp.
Location/Qualifiers

FEATURES

source

1. 438
/organism="Caenorhabditis elegans"
/mol_type="mRNA"
/strain="N2"
/db_xref="taxon:6239"
/clone="yk144912"
/sex="Hermaphrodite"
/tissue_type="whole animal"
/dev_stage="L4"
/clone_lib="unpublished oligo-capped cDNA library, stage I4"

ORIGIN

Query Match 85.6%; Score 15.4; DB 4; Length 438;
Best Local Similarity 94.1%; Pred. No. 5.5e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 AGACACACAGATTGCG 18
|||||
DB 142 AGACACACAGATTGCG 158

Search completed: December 3, 2004, 05:47:11
Job time: 2354.45 secs

This Page Blank (uspio)

OM nucleic - nucleic search, using sw model

Run on: December 2, 2004, 20:35:25 ; Search time 257.211 Seconds

(without alignments)
367.363 Million cell updates/sec

Title: US-10-050-189A-6

Perfect score: 18

Sequence: 1 gagaaacaagaatcgc 18

Scoring table:

IDENTITY_NUC

Gapop 10.0 ; Gapext 1.0

Searched: 4134886 seqs, 2624710521 residues

Total number of hits satisfying chosen parameters: 6269772

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 1000 summaries

Database :
1: Genesegm238604:*
2: Genesegm1980s:*
3: Genesegm1980s:*
4: Genesegm2000s:*
5: Genesegm2000s:*
6: Genesegm2000s:*
7: Genesegm2000s:*
8: Genesegm2000s:*
9: Genesegm2000s:*
10: Genesegm2000s:*
11: Genesegm2000s:*
12: Genesegm2000s:*Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result	No.	Score	Query	Match	Length	DB	ID	Description
C	1	16.4	91.1	65589	6	ABA90521.23		Confirmation (24 o
	2	16	88.9	19	6	ABN84784		ABN84784 Primer us
	3	16	88.9	66479	6	ABQ80567		ABQ80567 Mutant hu
	4	16	88.9	66479	6	ABQ80566		ABQ80566 Mutant hu
	5	16	88.9	66479	6	ABQ80568		ABQ80568 Mutant hu
	6	16	88.9	66479	6	ABQ80565		ABQ80565 Human IKK
	7	15.4	85.6	485	6	ABL37566		ABL37566 Human col
	8	15.4	85.6	513	8	ACA32902		ACA32902 Prokaryot
	9	15.4	85.6	534	10	ADH84638		ADH84638 Enterococ
	10	15.4	85.6	1470	4	AAS52481		AAS52481 E. coli D
	11	15.4	85.6	1822	10	ADH72695		ADH72695 Human end
	12	15.4	85.6	1944	5	AAS88539		AAS88539 DNA encod
	13	15.4	85.6	1944	5	AAS90079		AAS90079 DNA encod
	14	15.4	85.6	1944	5	AAS93115		AAS93115 DNA encod
	15	15.4	85.6	2186	4	AAH48474		AAH48474 Escherich
	16	15.4	85.6	2620	10	ADH72674		ADH72674 Human end
	17	15.4	85.6	2747	10	ADH72675		ADH72675 Human end
	18	15.4	85.6	2789	10	ADH72693		ADH72693 Human end
	19	15.4	85.6	3178	2	AAV30824		AAV30824 Mitochond
	20	15.4	85.6	3178	2	AAV42018		AAV42018 Glycero
	21	15.4	85.6	3178	2	AAV35739		AAV35739 Klebsiell

22	15.4	85.6	3178	2	AAH80613		AAH80613 CUT2 gene
23	15.4	85.6	3333	12	ADH72187		ADH72187 Human gen
24	15.4	85.6	3333	12	ADH72193		ADH72193 Human gen
25	15.4	85.6	3368	10	ADH72673		ADH72673 Human end
26	15.4	85.6	3470	10	ADH72686		ADH72686 Human end
27	15.4	85.6	3741	10	ADH72691		ADH72691 Human end
28	15.4	85.6	3858	10	ADH72694		ADH72694 Human end
29	15.4	85.6	3886	10	ADH72694		ADH72694 Human end
30	15.4	85.6	4067	10	ADH72697		ADH72697 Human end
31	15.4	85.6	4331	10	ADH72692		ADH72692 Human end
32	15.4	85.6	4404	10	ADH72671		ADH72671 Human end
33	15.4	85.6	4717	6	ABK99971		ABK99971 DNA encod
34	15.4	85.6	4751	10	ADH72672		ADH72672 Human end
35	15.4	85.6	4759	6	ABK99957		ABK99957 DNA encod
36	15.4	85.6	4762	8	AAH57275		AAH57275 LIG-1 'hu
37	15.4	85.6	4762	12	ADH72191		ADH72191 Human gen
38	15.4	85.6	5101	6	ABV93354		ABV93354 Human NOV
39	15.4	85.6	5101	12	ADH72189		ADH72189 Human gen
40	15.4	85.6	5188	10	ADH72682		ADH72682 Human end
41	15.4	85.6	5188	10	ADH72688		ADH72688 Human end
42	15.4	85.6	5204	10	ADH72689		ADH72689 Human end
43	15.4	85.6	5204	10	ADH72689		ADH72689 Human end
44	15.4	85.6	5204	10	ADH72689		ADH72689 Human end
45	15.4	85.6	5204	10	ADH72689		ADH72689 Human end
46	15.4	85.6	5204	10	ADH72689		ADH72689 Human end
47	15.4	85.6	5204	10	ADH72689		ADH72689 Human end
48	15.4	85.6	5204	10	ADH72689		ADH72689 Human end
49	15.4	85.6	5204	10	ADH72689		ADH72689 Human end
50	15.4	85.6	5204	10	ADH72689		ADH72689 Human end
51	15.4	85.6	5204	10	ADH72689		ADH72689 Human end
52	15.4	85.6	5204	10	ADH72689		ADH72689 Human end
53	15.4	85.6	5204	10	ADH72689		ADH72689 Human end
54	15.4	85.6	5204	10	ADH72689		ADH72689 Human end
55	15.4	85.6	5204	10	ADH72689		ADH72689 Human end
56	15.4	85.6	5204	10	ADH72689		ADH72689 Human end
57	15.4	85.6	5204	10	ADH72689		ADH72689 Human end
58	15.4	85.6	5204	10	ADH72689		ADH72689 Human end
59	15.4	85.6	5204	10	ADH72689		ADH72689 Human end
60	15.4	85.6	5204	10	ADH72689		ADH72689 Human end
61	15.4	85.6	5204	10	ADH72689		ADH72689 Human end
62	15.4	85.6	5204	10	ADH72689		ADH72689 Human end
63	15.4	85.6	5204	10	ADH72689		ADH72689 Human end
64	15.4	85.6	5204	10	ADH72689		ADH72689 Human end
65	15.4	85.6	5204	10	ADH72689		ADH72689 Human end
66	15.4	85.6	5204	10	ADH72689		ADH72689 Human end
67	15.4	85.6	5204	10	ADH72689		ADH72689 Human end
68	15.4	85.6	5204	10	ADH72689		ADH72689 Human end
69	15.4	85.6	5204	10	ADH72689		ADH72689 Human end
70	15.4	85.6	5204	10	ADH72689		ADH72689 Human end
71	15.4	85.6	5204	10	ADH72689		ADH72689 Human end
72	15.4	85.6	5204	10	ADH72689		ADH72689 Human end
73	15.4	85.6	5204	10	ADH72689		ADH72689 Human end
74	15.4	85.6	5204	10	ADH72689		ADH72689 Human end
75	15.4	85.6	5204	10	ADH72689		ADH72689 Human end
76	15.4	85.6	5204	10	ADH72689		ADH72689 Human end
77	15.4	85.6	5204	10	ADH72689		ADH72689 Human end
78	15.4	85.6	5204	10	ADH72689		ADH72689 Human end
79	15.4	85.6	5204	10	ADH72689		ADH72689 Human end
80	15.4	85.6	5204	10	ADH72689		ADH72689 Human end
81	15.4	85.6	5204	10	ADH72689		ADH72689 Human end
82	15.4	85.6	5204	10	ADH72689		ADH72689 Human end
83	15.4	85.6	5204	10	ADH72689		ADH72689 Human end
84	15.4	85.6	5204	10	ADH72689		ADH72689 Human end
85	15.4	85.6	5204	10	ADH72689		ADH72689 Human end
86	15.4	85.6	5204	10	ADH72689		ADH72689 Human end
87	15.4	85.6	5204	10	ADH72689		ADH72689 Human end
88	15.4	85.6	5204	10	ADH72689		ADH72689 Human end
89	15.4	85.6	5204	10	ADH72689		ADH72689 Human end
90	15.4	85.6	5204	10	ADH72689		ADH72689 Human end
91	15.4	85.6	5204	10	ADH72689		ADH72689 Human end
92	15.4	85.6	5204	10	ADH72689		ADH72689 Human end
93	15.4	85.6	5204	10	ADH72689		ADH72689 Human end
94	15.4	85.6	5204	10	ADH72689		ADH72689 Human end

95	14.8	82.2	2000	8	ADA73070	Adar73070 Rice gene	168	14.4	80.0	995	3	AAc77255	Aac47255 Arabidops
96	14.8	82.2	2157	4	ABL12755	Ab112755 Drosophill	169	14.4	80.0	1106	3	AAc54806	Aac54806 Arabidops
97	14.8	82.2	2488	8	ACA38431	Acac38431 Prokaryot	170	14.4	80.0	1109	3	AAc38856	Aac38856 Arabidops
98	14.8	82.2	2529	12	ADN72544	Adn72544 Thale cre	171	14.4	80.0	1159	3	AAAc7700	AAac7700 Human sec
99	14.8	82.2	2602	4	ABL12222	Ab112222 Drosophill	172	14.4	80.0	1159	8	ADA40321	Ada40321 Human sec
100	14.8	82.2	2712	8	ACA2040	Acac2040 Prokaryot	173	14.4	80.0	1159	10	ADAc56483	ADAc56483 Gene enco
101	14.8	82.2	2727	9	ADA31755	Ada31755 DNA enco	174	14.4	80.0	1200	3	AAAc67678	AAAc67678 Human sec
102	14.8	82.2	2821	4	ADAc5832	ADAc5832 Arabidops	175	14.4	80.0	1200	8	ADA39758	ADA39758 Human sec
103	14.8	82.2	2821	10	ADAc55633	ADAc55633 Thalecres	176	14.4	80.0	1200	10	ADAc55959	ADAc55959 Gene enco
104	14.8	82.2	2821	10	ADAc50000	ADAc50000 Plant yle	177	14.4	80.0	1357	5	AAAc6904	AAAc6904 DNA enco
105	14.8	82.2	2821	12	ADAc16164	ADAc16164 Plant tra	178	14.4	80.0	1413	11	ADAc7034	ADAc7034 Aspergill
106	14.8	82.2	2821	12	ADAc03362	ADAc03362 Thalecres	179	14.4	80.0	1413	11	ADAc7093	ADAc7093 Aspergill
107	14.8	82.2	2821	12	ADAc03362	ADAc03362 Thalecres	180	14.4	80.0	1413	11	ADAc7093	ADAc7093 Aspergill
108	14.8	82.2	2821	12	ADAc03362	ADAc03362 Thalecres	181	14.4	80.0	1413	11	ADAc7093	ADAc7093 Aspergill
109	14.8	82.2	2821	12	ADAc03362	ADAc03362 Thalecres	182	14.4	80.0	1413	11	ADAc7093	ADAc7093 Aspergill
110	14.8	82.2	3081	10	ADAc11353	ADAc11353 Drosophill	183	14.4	80.0	1452	2	AAAc75082	AAAc75082 Steaphyloc
111	14.8	82.2	3081	10	ADAc11353	ADAc11353 Drosophill	184	14.4	80.0	1722	6	AAAc75839	AAAc75839 Magnaport
112	14.8	82.2	3511	4	ABL11009	Ab111009 Drosophill	185	14.4	80.0	1820	4	AAAc60414	AAAc60414 Human pol
113	14.8	82.2	3511	10	ADAc11438	ADAc11438 Drosophill	186	14.4	80.0	1823	3	AAAc58628	AAAc58628 Human pol
114	14.8	82.2	3521	4	ABL12846	Ab112846 Drosophill	187	14.4	80.0	1823	5	ADAc98846	ADAc98846 DNA enco
115	14.8	82.2	3693	4	ABL120918	Ab1120918 Drosophill	188	14.4	80.0	1823	5	ADAc98846	ADAc98846 DNA enco
116	14.8	82.2	4197	4	AAAc07404	AAAc07404 Arabidops	189	14.4	80.0	1878	8	ADAc68528	ADAc68528 Arabidops
117	14.8	82.2	4197	4	AAAc07404	AAAc07404 Arabidops	190	14.4	80.0	1950	2	AAAc90402	AAAc90402 Nicotiana
118	14.8	82.2	4197	4	AAAc07404	AAAc07404 Arabidops	191	14.4	80.0	2177	12	ADAc889789	ADAc889789 Nicotiana
119	14.8	82.2	4197	4	AAAc07404	AAAc07404 Arabidops	192	14.4	80.0	2271	4	AAAc13358	AAAc13358 Human sec
120	14.8	82.2	4197	4	AAAc07404	AAAc07404 Arabidops	193	14.4	80.0	2359	10	ADAc62773	ADAc62773 Human ext
121	14.8	82.2	4197	4	AAAc07404	AAAc07404 Arabidops	194	14.4	80.0	2390	4	ADAc8046	ADAc8046 Human ext
122	14.8	82.2	4624	8	ACA45314	Acac45314 Prokaryot	195	14.4	80.0	2413	11	ADAc07092	ADAc07092 Aspergill
123	14.8	82.2	5631	4	ABL11758	Ab111758 Drosophill	196	14.4	80.0	2583	3	AAAc8752	AAAc8752 Arabidops
124	14.8	82.2	5824	4	ABL11008	Ab111008 Drosophill	197	14.4	80.0	2958	3	AAAc69470	AAAc69470 Human mal
125	14.8	82.2	6636	6	ABN80023	Abn80023 Human cre	198	14.4	80.0	3078	6	ABAc54988	ABAc54988 Human ova
126	14.8	82.2	7146	4	AAAc1616	AAAc1616 Drosophill	199	14.4	80.0	3125	6	ABAc54988	ABAc54988 Human ova
127	14.8	82.2	7146	4	AAAc1616	AAAc1616 Drosophill	200	14.4	80.0	3286	2	AAAc52267	AAAc52267 Protein p
128	14.8	82.2	7747	4	AAAc07401	AAAc07401 Arabidops	201	14.4	80.0	3286	2	AAAc58590	AAAc58590 Human PRO
129	14.8	82.2	8043	4	ABL02574	Ab102574 Drosophill	202	14.4	80.0	3286	3	ADAc78621	ADAc78621 Human PRO
130	14.8	82.2	8202	4	ABL09180	Ab109180 Drosophill	203	14.4	80.0	3286	4	AAAc72425	AAAc72425 Human PRO
131	14.8	82.2	8731	4	ABL11358	Ab111358 Drosophill	204	14.4	80.0	3286	4	AAAc1428	AAAc1428 Human CDN
132	14.8	82.2	35724	4	AAAc85022	AAAc85022 Adenovitu	205	14.4	80.0	3286	8	AAAc60256	AAAc60256 Human CDN
133	14.8	82.2	35724	4	AAAc85022	AAAc85022 Adenovitu	206	14.4	80.0	3286	8	AAAc07657	AAAc07657 Novel hum
134	14.8	82.2	44377	2	AAAc80414	AAAc80414 Platenoli	207	14.4	80.0	3286	8	AAAc03787	AAAc03787 CDNA enco
135	14.8	82.2	44377	2	AAAc80414	AAAc80414 Platenoli	208	14.4	80.0	3286	8	AAAc71705	AAAc71705 Human CDN
136	14.8	82.2	84428	12	ADAc45913	ADAc45913 Streptomy	209	14.4	80.0	3286	8	AAAc7036	AAAc7036 Human sec
137	14.8	82.2	110000	05	AAAc20248	AAAc20248 (6 of	210	14.4	80.0	3286	8	AAAc93235	AAAc93235 DNA enco
138	14.8	82.2	110000	05	AAAc20248	AAAc20248 (6 of	211	14.4	80.0	3286	8	AAAc93235	AAAc93235 DNA enco
139	14.8	82.2	110000	05	AAAc20248	AAAc20248 (6 of	212	14.4	80.0	3286	8	AAAc93235	AAAc93235 DNA enco
140	14.8	82.2	111309	2	AAAc20250	AAAc20250 Borrelia	213	14.4	80.0	3286	8	AAAc05594	AAAc05594 Human sec
141	14.4	80.0	174	12	AAAc86755	AAAc86755 Human gen	214	14.4	80.0	3286	8	AAAc04308	AAAc04308 Human CDN
142	14.4	80.0	174	12	AAAc86755	AAAc86755 Human gen	215	14.4	80.0	3286	8	AAAc55065	AAAc55065 Novel hum
143	14.4	80.0	369	4	AAAc69054	AAAc69054 Human INS	216	14.4	80.0	3286	9	ADAc45888	ADAc45888 Novel hum
144	14.4	80.0	439	4	AAAc71917	AAAc71917 Human cer	217	14.4	80.0	3286	9	ADAc76319	ADAc76319 Human PRO
145	14.4	80.0	439	4	AAAc71917	AAAc71917 Human cer	218	14.4	80.0	3286	9	ADAc29514	ADAc29514 Human sec
146	14.4	80.0	439	10	ADAc74705	ADAc74705 Colon can	219	14.4	80.0	3286	9	ADAc18969	ADAc18969 Human PRO
147	14.4	80.0	441	4	AAAc85921	AAAc85921 Human pol	220	14.4	80.0	3286	9	ADAc1592	ADAc1592 Homo sapi
148	14.4	80.0	478	9	AAAc15202	AAAc15202 Human adu	221	14.4	80.0	3286	9	ADAc19377	ADAc19377 Novel hum
149	14.4	80.0	493	9	AAAc44748	AAAc44748 Human fee	222	14.4	80.0	3286	9	ADAc27918	ADAc27918 CDNA enco
150	14.4	80.0	510	4	AAAc70717	AAAc70717 Human cer	223	14.4	80.0	3286	9	ADAc86397	ADAc86397 Novel hum
151	14.4	80.0	530	12	AAAc73064	AAAc73064 Human gen	224	14.4	80.0	3286	9	ADAc15961	ADAc15961 Human PRO
152	14.4	80.0	551	9	AAAc24266	AAAc24266 Human adu	225	14.4	80.0	3286	9	ADAc47747	ADAc47747 Human sec
153	14.4	80.0	551	9	AAAc24266	AAAc24266 Human adu	226	14.4	80.0	3286	9	ADAc18371	ADAc18371 Human sec
154	14.4	80.0	591	8	AAAc275080	AAAc275080 Mouse emb	227	14.4	80.0	3286	9	ADAc67546	ADAc67546 Human CDN
155	14.4	80.0	591	11	AAAc04328	AAAc04328 Pseudomon	228	14.4	80.0	3286	9	ADAc67546	ADAc67546 Human CDN
156	14.4	80.0	648	8	AAAc72703	AAAc72703 Steaphyloc	229	14.4	80.0	3286	9	ADAc30549	ADAc30549 CDNA enco
157	14.4	80.0	684	6	AAAc213158	AAAc213158 Arabidops	230	14.4	80.0	3286	9	ADAc85845	ADAc85845 Novel hum
158	14.4	80.0	684	6	AAAc213158	AAAc213158 Arabidops	231	14.4	80.0	3286	9	ADAc97057	ADAc97057 Human PRO
159	14.4	80.0	684	6	AAAc213158	AAAc213158 Arabidops	232	14.4	80.0	3286	9	ADAc97057	ADAc97057 Human PRO
160	14.4	80.0	684	6	AAAc213158	AAAc213158 Arabidops	233	14.4	80.0	3286	9	ADAc97057	ADAc97057 Human PRO
161	14.4	80.0	684	6	AAAc213158	AAAc213158 Arabidops	234	14.4	80.0	3286	9	ADAc97057	ADAc97057 Human PRO
162	14.4	80.0	701	3	AAAc03961	AAAc03961 Arabidops	235	14.4	80.0	3286	9	ADAc16702	ADAc16702 Human PRO
163	14.4	80.0	740	3	AAAc03961	AAAc03961 Arabidops	236	14.4	80.0	3286	9	ADAc83207	ADAc83207 Human PRO
164	14.4	80.0	783	11	ADAc01696	ADAc01696 Human col	237	14.4	80.0	3286	9	ADAc16346	ADAc16346 Human sec
165	14.4	80.0	808	8	AAAc72679	AAAc72679 Human cer	238	14.4	80.0	3286	9	ADAc19194	ADAc19194 Novel hum
166	14.4	80.0	867	8	AAAc73930	AAAc73930 Steaphyloc	239	14.4	80.0	3286	9	ADAc18818	ADAc18818 Human PRO
167	14.4	80.0	867	11	ADAc04519	ADAc04519 Pseudomon	240	14.4	80.0	3286	9	ADAc94033	ADAc94033 Human PRO

241	14.4	80.0	3296	9	ADB19929	Novel	hum	314	14.4	80.0	3296	10	ADC40927	Human	sec
242	14.4	80.0	3296	9	ADB13241	Human	PRO	315	14.4	80.0	3296	10	ADC19584	Human	sec
243	14.4	80.0	3296	9	ACD98608	Novel	hum	316	14.4	80.0	3296	10	ADC34032	Human	sec
244	14.4	80.0	3296	9	ADA74495	Human	PRO	317	14.4	80.0	3296	10	ADC13102	Human	sec
245	14.4	80.0	3296	9	ADA42491	Human	sec	318	14.4	80.0	3296	10	ADC50410	Novel	hum
246	14.4	80.0	3296	9	ADB24728	Human	PRO	319	14.4	80.0	3296	10	ADC71957	Novel	hum
247	14.4	80.0	3296	9	ADA82252	Human	PRO	320	14.4	80.0	3296	10	ADC59936	Novel	hum
248	14.4	80.0	3296	9	ADA75215	Human	PRO	321	14.4	80.0	3296	10	ADC52943	Novel	hum
249	14.4	80.0	3296	9	ADA85293	Novel	hum	322	14.4	80.0	3296	10	ADC57297	Novel	hum
250	14.4	80.0	3296	9	ADA84741	Novel	hum	323	14.4	80.0	3296	10	ADC60488	Novel	hum
251	14.4	80.0	3296	9	ACD23385	Human	PRO	324	14.4	80.0	3296	10	ADC50963	Novel	hum
252	14.4	80.0	3296	9	ADB23997	CDNA	encc	325	14.4	80.0	3296	10	ADC65490	Novel	hum
253	14.4	80.0	3296	9	ADA80525	Human	PRO	326	14.4	80.0	3296	10	ADC54588	Novel	hum
254	14.4	80.0	3296	9	ADA75767	Human	PRO	327	14.4	80.0	3296	10	ADC53549	Novel	hum
255	14.4	80.0	3296	9	ADA46992	Human	PRO	328	14.4	80.0	3296	10	ADC59072	Novel	hum
256	14.4	80.0	3296	9	ADB25288	Human	PRO	329	14.4	80.0	3296	10	ADC55950	Novel	hum
257	14.4	80.0	3296	9	ADA93464	Human	PRO	330	14.4	80.0	3296	10	ADC58520	Novel	hum
258	14.4	80.0	3296	9	ADB26814	CDNA	encc	331	14.4	80.0	3296	10	ADC12554	Novel	hum
259	14.4	80.0	3296	9	ADB31101	CDNA	encc	332	14.4	80.0	3296	10	ADC03194	Novel	hum
260	14.4	80.0	3296	9	ADA61029	Homo sapi		333	14.4	80.0	3296	10	ADC90186	Novel	hum
261	14.4	80.0	3296	9	ADB24176	Human	PRO	334	14.4	80.0	3296	10	ADC69605	Novel	hum
262	14.4	80.0	3296	9	ADA96505	Human	PRO	335	14.4	80.0	3296	10	ADC48494	Novel	hum
263	14.4	80.0	3296	9	ADA81077	Human	PRO	336	14.4	80.0	3296	10	ADD10023	Novel	hum
264	14.4	80.0	3296	9	ADA95953	Human	PRO	337	14.4	80.0	3296	10	ADD04598	Novel	hum
265	14.4	80.0	3296	9	ADB26262	CDNA	encc	338	14.4	80.0	3296	10	ADC80554	Novel	hum
266	14.4	80.0	3296	9	ADB21747	Novel	hum	339	14.4	80.0	3296	10	ADD11061	Human	PRO
267	14.4	80.0	3296	9	ADA77526	Human	PRO	340	14.4	80.0	3296	10	ADC47942	Novel	hum
268	14.4	80.0	3296	9	ADB18266	CDNA	encc	341	14.4	80.0	3296	10	ADD05109	Human	sec
269	14.4	80.0	3296	9	ADA86949	Novel	hum	342	14.4	80.0	3296	10	ADC80002	Novel	hum
270	14.4	80.0	3296	9	ADA16770	Human	sec	343	14.4	80.0	3296	10	ADD09471	Human	PRO
271	14.4	80.0	3296	9	ADA13199	Human	sec	344	14.4	80.0	3296	10	ADD04115	Novel	hum
272	14.4	80.0	3296	9	ADA42067	Human	sec	345	14.4	80.0	3296	10	ADD03691	Novel	hum
273	14.4	80.0	3296	9	ADA88052	Novel	hum	346	14.4	80.0	3296	10	ADD41184	Novel	hum
274	14.4	80.0	3296	9	ADA46440	Novel	hum	347	14.4	80.0	3296	10	ADD52323	CDNA	encc
275	14.4	80.0	3296	9	ADA17414	Human	sec	348	14.4	80.0	3296	10	ADD53063	Novel	hum
276	14.4	80.0	3296	9	ADA42917	Human	sec	349	14.4	80.0	3296	10	ADD53615	Novel	hum
277	14.4	80.0	3296	9	ADB28470	CDNA	encc	350	14.4	80.0	3296	10	ADD51771	Novel	hum
278	14.4	80.0	3296	9	ADB29022	CDNA	encc	351	14.4	80.0	3296	10	ADD02570	Human	PRO
279	14.4	80.0	3296	9	ADA76974	Human	PRO	352	14.4	80.0	3296	10	ADD02004	Human	PRO
280	14.4	80.0	3296	9	ADA88604	Novel	hum	353	14.4	80.0	3296	10	ADD54186	Novel	hum
281	14.4	80.0	3296	9	ADA97609	Human	PRO	354	14.4	80.0	3296	10	ADD92503	Human	PRO
282	14.4	80.0	3296	9	ADB27366	CDNA	encc	355	14.4	80.0	3296	10	ADD91399	Human	PRO
283	14.4	80.0	3296	9	ADB22299	Novel	hum	356	14.4	80.0	3296	10	ADD60403	Novel	hum
284	14.4	80.0	3296	9	ACD23747	Human	PRO	357	14.4	80.0	3296	10	ADB32310	Novel	hum
285	14.4	80.0	3296	9	ADA66990	Human	PRO	358	14.4	80.0	3296	10	ADB22242	CDNA	encc
286	14.4	80.0	3296	9	ADB22851	Human	PRO	359	14.4	80.0	3296	10	ADD79466	CDNA	encc
287	14.4	80.0	3296	9	ADB23624	Human	PRO	360	14.4	80.0	3296	10	ADB42002	Human	PRO
288	14.4	80.0	3296	9	ADA93246	Novel	hum	361	14.4	80.0	3296	10	ADB17819	Human	PRO
289	14.4	80.0	3296	9	ADB15409	Human	PRO	362	14.4	80.0	3296	10	ADD91951	Human	PRO
290	14.4	80.0	3296	9	ADB38661	Novel	hum	363	14.4	80.0	3296	10	ADB33414	Novel	hum
291	14.4	80.0	3296	9	ADB38109	Novel	hum	364	14.4	80.0	3296	10	ADB33866	Novel	hum
292	14.4	80.0	3296	9	ADB66581	Novel	hum	365	14.4	80.0	3296	10	ADD80018	Novel	hum
293	14.4	80.0	3296	10	ADB89661	Human	PRO	366	14.4	80.0	3296	10	ADD93055	Human	PRO
294	14.4	80.0	3296	10	ADB90393	Human	PRO	367	14.4	80.0	3296	10	ADB19475	Novel	hum
295	14.4	80.0	3296	10	ADB77835	Human	sec	368	14.4	80.0	3296	10	ADB34943	Novel	hum
296	14.4	80.0	3296	10	ADB39494	Novel	hum	369	14.4	80.0	3296	10	ADB18923	Human	PRO
297	14.4	80.0	3296	10	ADB74971	Human	sec	370	14.4	80.0	3296	10	ADB43119	Novel	hum
298	14.4	80.0	3296	10	ADB84717	Novel	hum	371	14.4	80.0	3296	10	ADD95908	Human	PRO
299	14.4	80.0	3296	10	ADB86724	Human	PRO	372	14.4	80.0	3296	10	ADB22794	Novel	hum
300	14.4	80.0	3296	10	ADB77329	Novel	hum	373	14.4	80.0	3296	10	ADB78912	Novel	hum
301	14.4	80.0	3296	10	ADB34486	Human	PRO	374	14.4	80.0	3296	10	ADB32862	Novel	hum
302	14.4	80.0	3296	10	ADB35590	Human	PRO	375	14.4	80.0	3296	10	ADB42554	Human	PRO
303	14.4	80.0	3296	10	ADB33934	Human	PRO	376	14.4	80.0	3296	10	ADB80570	CDNA	encc
304	14.4	80.0	3296	10	ADB35038	Human	PRO	377	14.4	80.0	3296	10	ADB89598	Human	PRO
305	14.4	80.0	3296	10	ADB36142	Human	PRO	378	14.4	80.0	3296	10	ADB40882	Novel	hum
306	14.4	80.0	3296	10	ADB46537	Novel	hum	379	14.4	80.0	3296	10	ADB50481	Novel	hum
307	14.4	80.0	3296	10	ADC28618	Human	sec	380	14.4	80.0	3296	10	ADB92810	Novel	hum
308	14.4	80.0	3296	10	ADC39818	Human	sec	381	14.4	80.0	3296	10	ADG21519	Novel	hum
309	14.4	80.0	3296	10	ADC40332	Human	sec	382	14.4	80.0	3296	10	ADG23160	Novel	hum
310	14.4	80.0	3296	10	ADC19156	Human	sec	383	14.4	80.0	3296	10	ADG97495	Human	PRO
311	14.4	80.0	3296	10	ADC34456	Human	sec	384	14.4	80.0	3296	10	ADG80559	Human	PRO
312	14.4	80.0	3296	10	ADC29511	Human	sec	385	14.4	80.0	3296	10	ADG80007	Human	PRO
313	14.4	80.0	3296	10	ADC29042	Human	sec	386	14.4	80.0	3296	10	ADH59426	Human	sec

387	14.4	80.0	3296	10	ADH55299	Adh55299	Novel	hum	460	14.4	80.0	3296	12	ADG82215	Adg82215	Human	PRO
388	14.4	80.0	3296	10	ADH55851	Adh55851	Novel	hum	461	14.4	80.0	3296	12	ADG57454	Adg57454	Novel	hum
389	14.4	80.0	3296	10	ADH38205	Adh38205	Human	sec	462	14.4	80.0	3296	12	ADG56902	Adg56902	Novel	hum
390	14.4	80.0	3296	10	ADH64070	Adh64070	Novel	hum	463	14.4	80.0	3296	12	ADG55798	Adg55798	Novel	hum
391	14.4	80.0	3296	10	ADH65019	Adh65019	Novel	hum	464	14.4	80.0	3296	12	ADG58558	Adg58558	Novel	hum
392	14.4	80.0	3296	10	ADH63518	Adh63518	Novel	hum	465	14.4	80.0	3296	12	ADG70924	Adg70924	Novel	hum
393	14.4	80.0	3296	10	ADH81932	Adh81932	Novel	hum	466	14.4	80.0	3296	12	ADG92833	Adg92833	Human	sec
394	14.4	80.0	3296	10	ADH81380	Adh81380	Novel	hum	467	14.4	80.0	3296	12	ADG58006	Adg58006	Novel	hum
395	14.4	80.0	3296	10	ACAS59152	Acas59152	Human	PRO	468	14.4	80.0	3296	12	ADG53590	Adg53590	Novel	hum
396	14.4	80.0	3296	10	ACD24037	Acad24037	Novel	hum	469	14.4	80.0	3296	12	ADG71476	Adg71476	Novel	hum
397	14.4	80.0	3296	10	ACAS58549	Acas58549	CDNA	encc	470	14.4	80.0	3296	12	ADG81663	Adg81663	Human	PRO
398	14.4	80.0	3296	10	ACA67178	Acac67178	CDNA	encc	471	14.4	80.0	3296	12	ADH30625	Adh30625	Human	PRO
399	14.4	80.0	3296	10	ADJ26473	Adj26473	Human	sec	472	14.4	80.0	3296	12	ADH11992	Adh11992	Novel	hum
400	14.4	80.0	3296	11	ADH62549	Adhm62549	Novel	hum	473	14.4	80.0	3296	12	ADG52414	Adg52414	Novel	hum
401	14.4	80.0	3296	11	ADN15948	Adn15948	Novel	hum	474	14.4	80.0	3296	12	ADG54142	Adg54142	Novel	hum
402	14.4	80.0	3296	11	ADN16577	Adn16577	Novel	hum	475	14.4	80.0	3296	12	ADG81111	Adg81111	Human	PRO
403	14.4	80.0	3296	11	ADN15396	Adn15396	Novel	hum	476	14.4	80.0	3296	12	ADG56350	Adg56350	Novel	hum
404	14.4	80.0	3296	11	ADN14844	Adn14844	Novel	hum	477	14.4	80.0	3296	12	ADH12616	Adh12616	Novel	hum
405	14.4	80.0	3296	12	ADCG1106	Adcg1106	Novel	hum	478	14.4	80.0	3296	12	ADG61462	Adg61462	Novel	hum
406	14.4	80.0	3296	12	ADH79388	Adh79388	Human	sec	479	14.4	80.0	3296	12	ADH28549	Adh28549	Human	PRO
407	14.4	80.0	3296	12	ADH76554	Adh76554	Human	PRO	480	14.4	80.0	3296	12	ADG54694	Adg54694	Novel	hum
408	14.4	80.0	3296	12	ADH87918	Adh87918	Human	PRO	481	14.4	80.0	3296	12	ADG59734	Adg59734	Novel	hum
409	14.4	80.0	3296	12	ADH63222	Adh63222	Human	PRO	482	14.4	80.0	3296	12	ADH20622	Adh20622	Human	sec
410	14.4	80.0	3296	12	ADH79812	Adh79812	Human	sec	483	14.4	80.0	3296	12	ADH07477	Adh07477	Human	sec
411	14.4	80.0	3296	12	ADH75770	Adh75770	Human	PRO	484	14.4	80.0	3296	12	ADH60022	Adh60022	Human	sec
412	14.4	80.0	3296	12	ADH73488	Adh73488	Human	sec	485	14.4	80.0	3296	12	ADH07050	Adh07050	Human	sec
413	14.4	80.0	3296	12	ADH23346	Adh23346	CDNA	encc	486	14.4	80.0	3296	12	ADH18158	Adh18158	CDNA	encc
414	14.4	80.0	3296	12	ADH23898	Adh23898	CDNA	encc	487	14.4	80.0	3296	12	ADH18792	Adh18792	Human	sec
415	14.4	80.0	3296	12	ADH24541	Adh24541	CDNA	encc	488	14.4	80.0	3296	12	ADH16512	Adh16512	Human	sec
416	14.4	80.0	3296	12	ADH87366	Adh87366	Human	PRO	489	14.4	80.0	3296	12	ADH137771	Adh137771	Human	sec
417	14.4	80.0	3296	12	ADH89232	Adh89232	Human	PRO	490	14.4	80.0	3296	12	ADG09901	Adg09901	Novel	hum
418	14.4	80.0	3296	12	ADH74023	Adh74023	Human	sec	491	14.4	80.0	3296	12	ADH97571	Adh97571	Human	sec
419	14.4	80.0	3296	12	ADH18371	Adh18371	Human	PRO	492	14.4	80.0	3296	12	ADH15372	Adh15372	Novel	hum
420	14.4	80.0	3296	12	ADH88680	Adh88680	Human	PRO	493	14.4	80.0	3296	12	ADG02449	Adg02449	Novel	hum
421	14.4	80.0	3296	12	ADH95977	Adh95977	Human	sec	494	14.4	80.0	3296	12	ADH165939	Adh165939	Human	sec
422	14.4	80.0	3296	12	ADH94700	Adh94700	CDNA	encc	495	14.4	80.0	3296	12	ADH14704	Adh14704	Novel	hum
423	14.4	80.0	3296	12	ADH91111	Adh91111	Human	PRO	496	14.4	80.0	3296	12	ADH60682	Adh60682	Human	sec
424	14.4	80.0	3296	12	ADH95252	Adh95252	CDNA	encc	497	14.4	80.0	3296	12	ADH18299	Adh18299	Novel	hum
425	14.4	80.0	3296	12	ADH93362	Adh93362	Human	PRO	498	14.4	80.0	3296	12	ADH97339	Adh97339	Human	sec
426	14.4	80.0	3296	12	ADH34943	Adh34943	CDNA	encc	499	14.4	80.0	3296	12	ADH08932	Adh08932	Human	sec
427	14.4	80.0	3296	12	ADH38686	Adh38686	Human	sec	500	14.4	80.0	3296	12	ADH25273	Adh25273	Human	sec
428	14.4	80.0	3296	12	ADH92258	Adh92258	Novel	hum	501	14.4	80.0	3296	12	ADH63580	Adh63580	Novel	hum
429	14.4	80.0	3296	12	ADH90559	Adh90559	Human	PRO	502	14.4	80.0	3296	12	ADH30023	Adh30023	Human	sec
430	14.4	80.0	3296	12	ADH91706	Adh91706	Novel	hum	503	14.4	80.0	3296	12	ADH77475	Adh77475	Human	PRO
431	14.4	80.0	3296	12	ADH99123	Adh99123	Human	sec	504	14.4	80.0	3296	12	ADH65597	Adh65597	CDNA	encc
432	14.4	80.0	3296	12	ADH40593	Adh40593	Human	sec	505	14.4	80.0	3296	12	ADH27733	Adh27733	CDNA	encc
433	14.4	80.0	3296	12	ADH73987	Adh73987	Human	sec	506	14.4	80.0	3296	12	ADH42457	Adh42457	CDNA	encc
434	14.4	80.0	3296	12	ADH02285	Adh02285	Human	PRO	507	14.4	80.0	3296	12	ADH06345	Adh06345	Human	PRO
435	14.4	80.0	3296	12	ADH22071	Adh22071	Novel	hum	508	14.4	80.0	3296	12	ADH28319	Adh28319	CDNA	encc
436	14.4	80.0	3296	12	ADH20141	Adh20141	CDNA	encc	509	14.4	80.0	3296	12	AAH69902	Aah69902	DNA	encc
437	14.4	80.0	3296	12	ADH98047	Adh98047	Human	PRO	510	14.4	80.0	4140	5	AAH69903	Aah69903	DNA	encc
438	14.4	80.0	3296	12	ADH24264	Adh24264	Novel	hum	511	14.4	80.0	4338	12	ADH026800	Adh026800	Human	INS
439	14.4	80.0	3296	12	ADH98618	Adh98618	Human	PRO	512	14.4	80.0	4426	8	ABX70463	Abx70463	DNA	encc
440	14.4	80.0	3296	12	ADH03449	Adh03449	Human	PRO	513	14.4	80.0	4426	12	ADH72209	Adh72209	Human	gen
441	14.4	80.0	3296	12	ADH99170	Adh99170	Human	PRO	514	14.4	80.0	4426	12	ADH24078	Adh24078	Human	NOV
442	14.4	80.0	3296	12	ADH16755	Adh16755	CDNA	encc	515	14.4	80.0	4488	2	AAV74539	Aav74539	Staphyloc	
443	14.4	80.0	3296	12	ADH05214	Adh05214	Human	PRO	516	14.4	80.0	4501	8	ABX70464	Abx70464	DNA	encc
444	14.4	80.0	3296	12	ADH19481	Adh19481	CDNA	encc	517	14.4	80.0	4501	12	ADH72197	Adh72197	Human	gen
445	14.4	80.0	3296	12	ADH73563	Adh73563	CDNA	encc	518	14.4	80.0	4501	12	ADL24080	Adl24080	Human	NOV
446	14.4	80.0	3296	12	ADH13318	Adh13318	CDNA	encc	519	14.4	80.0	4724	12	ADH06810	Adh06810	Human	INS
447	14.4	80.0	3296	12	ADH08375	Adh08375	Novel	hum	520	14.4	80.0	4940	2	AAV21451	Aav21451	P. faelcip	
448	14.4	80.0	3296	12	ADH15845	Adh15845	CDNA	encc	521	14.4	80.0	4994	12	ADH72195	Adh72195	Human	gen
449	14.4	80.0	3296	12	ADH96943	Adh96943	Human	PRO	522	14.4	80.0	5425	12	ADN05323	Adn05323	Antipsoit	
450	14.4	80.0	3296	12	ADH06128	Adh06128	Human	PRO	523	14.4	80.0	6778	4	ABL03470	AbL03470	Drosophill	
451	14.4	80.0	3296	12	ADH23712	Adh23712	Novel	hum	524	14.4	80.0	7401	4	ABL19174	AbL19174	Drosophill	
452	14.4	80.0	3296	12	ADH04001	Adh04001	Human	PRO	525	14.4	80.0	12123	4	AAK90226	Aak90226	Human	dig
453	14.4	80.0	3296	12	ADH24902	Adh24902	Novel	hum	526	14.4	80.0	12123	5	AAK39855	Aak39855	Gemomic s	
454	14.4	80.0	3296	12	ADH07199	Adh07199	Novel	hum	527	14.4	80.0	12123	9	ADB32815	Adb32815	Human	nov
455	14.4	80.0	3296	12	ADH07751	Adh07751	Novel	hum	528	14.4	80.0	24934	4	AAK78433	Aak78433	Human	imm
456	14.4	80.0	3296	12	ADH55246	Adh55246	Novel	hum	529	14.4	80.0	24934	4	AAK71542	Aak71542	Human	imm
457	14.4	80.0	3296	12	ADH60910	Adh60910	Novel	hum	530	14.4	80.0	110000	3	AAAF22303_0	Aaf22303_0	Atreidops	
458	14.4	80.0	3296	12	ADH62014	Adh62014	Novel	hum	531	14.4	80.0	110000	3	AAAF22303_29	Aaf22303_29	Continuation	(30 o
459	14.4	80.0	3296	12	ADH92406	Adh92406	Human	sec	532	14.4	80.0	110000	4	AAI99682_30	Aai99682_30	Continuation	(31 o

C 533	14.4	80.0	110000	4	AA199683_29	Continuation (30 o	605	13.8	76.7	267	6	AAD43726	Aad43726 Human TR1
534	14.4	80.0	110000	5	AA161375_0	AA161375 Soybean 3	607	13.8	76.7	267	9	ACH66634	ACH66634 Human TR1
C 535	14.4	80.0	110000	6	ABA90521_11	Continuation (12 o	608	13.8	76.7	267	9	ACH66195	ACH66195 Human TR1
C 536	14.4	80.0	110000	6	ABA90521_12	Continuation (13 o	609	13.8	76.7	273	10	ABX84860	ABX84860 Human TR1
C 537	14.4	80.0	110000	6	ABA90521_13	Continuation (14 o	610	13.8	76.7	288	8	ABX54929	ABX54929 Bovine ES
C 538	14.4	80.0	110000	8	ABQ83210_0	Abq83210 Human tra	611	13.8	76.7	289	10	ABX84297	ABX84297 Corn ear-
C 539	14.4	80.0	119596	6	ABQ83210_10	Abq83210 Human G-P	612	13.8	76.7	301	10	ABX85915	ABX85915 Corn ear-
540	14.4	80.0	162450	2	AAQ285967	AAQ285967 Retinobla	613	13.8	76.7	305	12	ACH86852	ACH86852 Human TR1
C 541	14.4	80.0	162450	2	AAQ285967	AAQ285967 Retinobla	614	13.8	76.7	315	8	ABX52904	ABX52904 Human TR1
C 542	14.4	80.0	162450	2	AAQ285967	AAQ285967 Retinobla	615	13.8	76.7	322	8	ACH14949	ACH14949 Human TR1
C 543	14.4	80.0	162450	2	AAQ285967	AAQ285967 Retinobla	616	13.8	76.7	333	10	ADQ91578	ADQ91578 Human TR1
544	14.4	80.0	162450	2	AAQ285967	AAQ285967 Retinobla	617	13.8	76.7	336	4	AAQ51598	AAQ51598 Human TR1
545	14.4	80.0	162450	2	AAQ285967	AAQ285967 Retinobla	618	13.8	76.7	351	8	ACA46492	ACA46492 Human TR1
C 546	14.4	80.0	162450	2	AAQ285967	AAQ285967 Retinobla	619	13.8	76.7	351	8	ACF74712	ACF74712 Human TR1
C 547	14.4	80.0	162450	2	AAQ285967	AAQ285967 Retinobla	620	13.8	76.7	354	4	AAQ55030	AAQ55030 Human TR1
548	14.4	80.0	162450	2	AAQ285967	AAQ285967 Retinobla	621	13.8	76.7	354	4	AAQ55452	AAQ55452 Human TR1
549	14.4	80.0	162450	2	AAQ285967	AAQ285967 Retinobla	622	13.8	76.7	354	4	ACA20061	ACA20061 Human TR1
550	14.4	80.0	162450	2	AAQ285967	AAQ285967 Retinobla	623	13.8	76.7	372	6	ABN90983	ABN90983 Human TR1
551	14.4	80.0	162450	2	AAQ285967	AAQ285967 Retinobla	624	13.8	76.7	391	3	AAQ75788	AAQ75788 Human TR1
552	14.4	80.0	162450	2	AAQ285967	AAQ285967 Retinobla	625	13.8	76.7	411	4	AAQ60155	AAQ60155 Human TR1
C 553	14.4	80.0	162450	2	AAQ285967	AAQ285967 Retinobla	626	13.8	76.7	456	4	ABL28221	ABL28221 Human TR1
C 554	14.4	80.0	162450	2	AAQ285967	AAQ285967 Retinobla	627	13.8	76.7	462	2	ABN99594	ABN99594 Human TR1
C 555	14.4	80.0	162450	2	AAQ285967	AAQ285967 Retinobla	628	13.8	76.7	462	2	ABN99594	ABN99594 Human TR1
C 556	14.4	80.0	162450	2	AAQ285967	AAQ285967 Retinobla	629	13.8	76.7	465	4	AAI34016	AAI34016 Human TR1
C 557	14.4	80.0	162450	2	AAQ285967	AAQ285967 Retinobla	630	13.8	76.7	465	4	ABA24131	ABA24131 Human TR1
C 558	14.4	80.0	162450	2	AAQ285967	AAQ285967 Retinobla	631	13.8	76.7	465	4	ABK28087	ABK28087 Human TR1
C 559	14.4	80.0	162450	2	AAQ285967	AAQ285967 Retinobla	632	13.8	76.7	465	4	ABK28087	ABK28087 Human TR1
C 560	14.4	80.0	162450	2	AAQ285967	AAQ285967 Retinobla	633	13.8	76.7	465	5	AAI02574	AAI02574 Human TR1
C 561	14.4	80.0	162450	2	AAQ285967	AAQ285967 Retinobla	634	13.8	76.7	465	5	ABK02581	ABK02581 Human TR1
C 562	14.4	80.0	162450	2	AAQ285967	AAQ285967 Retinobla	635	13.8	76.7	477	2	AAQ40654	AAQ40654 Human TR1
C 563	14.4	80.0	162450	2	AAQ285967	AAQ285967 Retinobla	636	13.8	76.7	478	6	AAQ99070	AAQ99070 Human TR1
C 564	14.4	80.0	162450	2	AAQ285967	AAQ285967 Retinobla	637	13.8	76.7	480	12	ADU44147	ADU44147 Human TR1
C 565	14.4	80.0	162450	2	AAQ285967	AAQ285967 Retinobla	638	13.8	76.7	482	3	AAQ38133	AAQ38133 Human TR1
566	14.4	80.0	162450	2	AAQ285967	AAQ285967 Retinobla	639	13.8	76.7	483	4	AAQ53875	AAQ53875 Human TR1
567	14.4	80.0	162450	2	AAQ285967	AAQ285967 Retinobla	640	13.8	76.7	486	10	ADK53676	ADK53676 Human TR1
568	14.4	80.0	162450	2	AAQ285967	AAQ285967 Retinobla	641	13.8	76.7	486	10	ADU40471	ADU40471 Human TR1
C 569	14.4	80.0	162450	2	AAQ285967	AAQ285967 Retinobla	642	13.8	76.7	488	3	AAQ53552	AAQ53552 Human TR1
C 570	14.4	80.0	162450	2	AAQ285967	AAQ285967 Retinobla	643	13.8	76.7	510	8	ABK53326	ABK53326 Human TR1
571	14.4	80.0	162450	2	AAQ285967	AAQ285967 Retinobla	644	13.8	76.7	512	12	ACH73121	ACH73121 Human TR1
C 572	14.4	80.0	162450	2	AAQ285967	AAQ285967 Retinobla	645	13.8	76.7	520	4	AAI16066	AAI16066 Human TR1
573	14.4	80.0	162450	2	AAQ285967	AAQ285967 Retinobla	646	13.8	76.7	531	10	ABK57470	ABK57470 Human TR1
574	14.4	80.0	162450	2	AAQ285967	AAQ285967 Retinobla	647	13.8	76.7	543	11	ABD03624	ABD03624 Human TR1
575	14.4	80.0	162450	2	AAQ285967	AAQ285967 Retinobla	648	13.8	76.7	546	2	AAQ99661	AAQ99661 Human TR1
576	14.4	80.0	162450	2	AAQ285967	AAQ285967 Retinobla	649	13.8	76.7	548	2	AAI184270	AAI184270 Human TR1
577	14.4	80.0	162450	2	AAQ285967	AAQ285967 Retinobla	650	13.8	76.7	558	6	ABA50401	ABA50401 Human TR1
C 578	14.4	80.0	162450	2	AAQ285967	AAQ285967 Retinobla	651	13.8	76.7	580	6	ABV02684	ABV02684 Human TR1
579	14.4	80.0	162450	2	AAQ285967	AAQ285967 Retinobla	652	13.8	76.7	641	5	ABQ29568	ABQ29568 Human TR1
C 580	14.4	80.0	162450	2	AAQ285967	AAQ285967 Retinobla	653	13.8	76.7	645	6	ABQ29568	ABQ29568 Human TR1
C 581	14.4	80.0	162450	2	AAQ285967	AAQ285967 Retinobla	654	13.8	76.7	645	6	ABQ29568	ABQ29568 Human TR1
C 582	14.4	80.0	162450	2	AAQ285967	AAQ285967 Retinobla	655	13.8	76.7	650	10	ADK59898	ADK59898 Human TR1
C 583	14.4	80.0	162450	2	AAQ285967	AAQ285967 Retinobla	656	13.8	76.7	653	3	ABK74457	ABK74457 Human TR1
584	14.4	80.0	162450	2	AAQ285967	AAQ285967 Retinobla	657	13.8	76.7	678	3	AAI13659	AAI13659 Human TR1
585	14.4	80.0	162450	2	AAQ285967	AAQ285967 Retinobla	658	13.8	76.7	686	6	ABQ124910	ABQ124910 Human TR1
586	14.4	80.0	162450	2	AAQ285967	AAQ285967 Retinobla	659	13.8	76.7	696	6	ABQ19481	ABQ19481 Human TR1
587	14.4	80.0	162450	2	AAQ285967	AAQ285967 Retinobla	660	13.8	76.7	696	6	ABQ19481	ABQ19481 Human TR1
C 588	14.4	80.0	162450	2	AAQ285967	AAQ285967 Retinobla	661	13.8	76.7	705	6	ABK74533	ABK74533 Human TR1
589	14.4	80.0	162450	2	AAQ285967	AAQ285967 Retinobla	662	13.8	76.7	712	4	AAH70542	AAH70542 Human TR1
590	14.4	80.0	162450	2	AAQ285967	AAQ285967 Retinobla	663	13.8	76.7	725	9	ADH14447	ADH14447 Human TR1
591	14.4	80.0	162450	2	AAQ285967	AAQ285967 Retinobla	664	13.8	76.7	728	9	ADH83031	ADH83031 Human TR1
592	14.4	80.0	162450	2	AAQ285967	AAQ285967 Retinobla	665	13.8	76.7	769	4	AAI37344	AAI37344 Human TR1
593	14.4	80.0	162450	2	AAQ285967	AAQ285967 Retinobla	666	13.8	76.7	769	4	ABK60332	ABK60332 Human TR1
594	14.4	80.0	162450	2	AAQ285967	AAQ285967 Retinobla	667	13.8	76.7	769	12	ADJ31082	ADJ31082 Human TR1
595	14.4	80.0	162450	2	AAQ285967	AAQ285967 Retinobla	668	13.8	76.7	778	6	ABQ89992	ABQ89992 Human TR1
596	14.4	80.0	162450	2	AAQ285967	AAQ285967 Retinobla	669	13.8	76.7	783	4	AAQ54375	AAQ54375 Human TR1
C 597	14.4	80.0	162450	2	AAQ285967	AAQ285967 Retinobla	670	13.8	76.7	783	4	AAQ442790	AAQ442790 Human TR1
C 598	14.4	80.0	162450	2	AAQ285967	AAQ285967 Retinobla	671	13.8	76.7	801	4	AAQ57018	AAQ57018 Human TR1
C 599	14.4	80.0	162450	2	AAQ285967	AAQ285967 Retinobla	672	13.8	76.7	801	10	ADD42812	ADD42812 Human TR1
600	14.4	80.0	162450	2	AAQ285967	AAQ285967 Retinobla	673	13.8	76.7	846	8	ACA18772	ACA18772 Human TR1
601	14.4	80.0	162450	2	AAQ285967	AAQ285967 Retinobla	674	13.8	76.7	849	8	ACA48818	ACA48818 Human TR1
C 602	14.4	80.0	162450	2	AAQ285967	AAQ285967 Retinobla	675	13.8	76.7	849	12	ADO05534	ADO05534 Human TR1
603	14.4	80.0	162450	2	AAQ285967	AAQ285967 Retinobla	676	13.8	76.7	852	8	ACA51219	ACA51219 Human TR1
604	14.4	80.0	162450	2	AAQ285967	AAQ285967 Retinobla	677	13.8	76.7	852	8	ACA52099	ACA52099 Human TR1
605	14.4	80.0	162450	2	AAQ285967	AAQ285967 Retinobla	678	13.8	76.7	852	10	ADD45795	ADD45795 Human TR1

679	13.8	76.7	869	2	AAX20199	Aax20199	Enterococ	C	752	13.8	76.7	1632	11	ABD13700	Abd13700	Pseudomon
680	13.8	76.7	869	6	ABN98184	Abn98184	E. faecali	C	753	13.8	76.7	1677	10	ADH84959	Adh84959	Enterococ
681	13.8	76.7	869	8	ACA88148	Aca88148	E. faecali	C	754	13.8	76.7	1680	2	AAX20042	Aax20042	Enterococ
682	13.8	76.7	869	8	ABX61754	Abx61754	Enterococ	C	755	13.8	76.7	1680	6	ABN98027	Abn98027	E. faecali
683	13.8	76.7	876	8	ACA42078	Aca42078	Prokaryot	C	756	13.8	76.7	1680	8	ACA87991	ACA87991	E. faecali
684	13.8	76.7	876	8	ABZ15661	Abz15661	Arabidops	C	757	13.8	76.7	1680	8	AXK61597	AXK61597	Enterococ
685	13.8	76.7	897	6	ADA70428	Ada70428	Rice gene	C	758	13.8	76.7	1680	11	AD061534	Ado61534	Transcrip
686	13.8	76.7	913	5	AAH81625	Aah81625	Human dif	C	759	13.8	76.7	1695	11	ABD13590	Abd13590	Pseudomon
687	13.8	76.7	917	3	AAC45004	Aac45004	Arabidops	C	760	13.8	76.7	1696	10	ADC03423	Adc03423	Rice flow
688	13.8	76.7	921	11	ABD03698	Abd03698	Pseudomon	C	761	13.8	76.7	1717	2	AAZ41991	Aaz41991	Human end
689	13.8	76.7	924	4	AAS56173	Aas56173	Salmonell	C	762	13.8	76.7	1719	6	AAZ20147	Aaz20147	Parainflu
690	13.8	76.7	924	4	AAI60687	Aai60687	Human pol	C	763	13.8	76.7	1728	10	ADFA2499	Adfa2499	Chinese h
691	13.8	76.7	933	6	ABX80162	Abx80162	Bacillus	C	764	13.8	76.7	1739	8	ABT20694	Abt20694	Aspergill
692	13.8	76.7	951	5	AAH68382	Aah68382	C. glutam	C	765	13.8	76.7	1739	8	ABT18678	Abt18678	Aspergill
693	13.8	76.7	951	5	ACA01298	Aca01298	C. glutam	C	766	13.8	76.7	1749	6	ABZ15564	Abz15564	Arabidops
694	13.8	76.7	951	10	ADG88469	Adg88469	Arabidops	C	767	13.8	76.7	1758	4	ABL21841	AbL21841	Drosophi
695	13.8	76.7	959	6	ABO38182	Abg38182	Oligonuci	C	768	13.8	76.7	1762	10	ADP82372	Adp82372	Leukemia
696	13.8	76.7	959	6	ABO38183	Abg38183	Oligonuci	C	769	13.8	76.7	1807	4	AAK52630	Aak52630	Human pol
697	13.8	76.7	984	10	ADC93261	Adc93261	E. faecali	C	770	13.8	76.7	1819	10	ADB94879	Adb94879	Programme
698	13.8	76.7	984	4	AAK51646	Aak51646	Human pol	C	771	13.8	76.7	1828	4	AAK526011	Aak526011	Human CDN
699	13.8	76.7	999	10	ADH84371	Adh84371	Enterococ	C	772	13.8	76.7	1828	8	ABX73352	Abx73352	Human nov
700	13.8	76.7	1001	3	AAC40948	Aac40948	Arabidops	C	773	13.8	76.7	1854	4	AAI158901	Aai158901	Human pol
701	13.8	76.7	1002	2	AAX20198	Aax20198	Enterococ	C	774	13.8	76.7	1854	5	ADQ99123	Adq99123	DNA encod
702	13.8	76.7	1002	6	ABN98183	Abn98183	E. faecali	C	775	13.8	76.7	1854	6	ADB48883	Adb48883	Novel hum
703	13.8	76.7	1002	8	ACA88147	Aca88147	E. faecali	C	776	13.8	76.7	1874	6	ABK35797	Abk35797	CDNA sequ
704	13.8	76.7	1002	8	ABX61753	Abx61753	Enterococ	C	777	13.8	76.7	1883	12	AD143296	Adi43296	Plant tra
705	13.8	76.7	1006	3	AAC40943	Aac40943	Arabidops	C	778	13.8	76.7	1888	2	AAQ05319	Aaq05319	Viral seq
706	13.8	76.7	1014	12	AD030222	Ado30222	Mouse GPC	C	779	13.8	76.7	1930	5	AAK92774	Aak92774	DNA encod
707	13.8	76.7	1017	6	ABO70253	Abg70253	Listeria	C	780	13.8	76.7	1949	10	ADC91379	Adc91379	E. faecali
708	13.8	76.7	1041	11	ABD04006	Abd04006	Pseudomon	C	781	13.8	76.7	1955	6	ABO99309	AbO99309	Human cod
709	13.8	76.7	1080	3	AAC53690	Aac53690	Arabidops	C	782	13.8	76.7	1965	8	ACA19502	Aca19502	Prokaryot
710	13.8	76.7	1124	2	AAQ82820	Aaq82820	A. aculea	C	783	13.8	76.7	2000	6	ABZ17042	Abz17042	Arabidops
711	13.8	76.7	1125	6	ABZ12421	Abz12421	Arabidops	C	784	13.8	76.7	2045	11	ABD13656	Abd13656	Pseudomon
712	13.8	76.7	1133	2	AAZ25118	Aaz25118	Corn anth	C	785	13.8	76.7	2055	2	AAT38260	Aat38260	10 gene d
713	13.8	76.7	1149	8	ACF33366	Acf33366	Mycobacte	C	786	13.8	76.7	2055	3	AAK51893	Aak51893	T helper
714	13.8	76.7	1149	8	ACA38749	Aca38749	Prokaryot	C	787	13.8	76.7	2055	4	AAC90981	Aac90981	Mouse ful
715	13.8	76.7	1152	8	ACA40811	Aca40811	Prokaryot	C	788	13.8	76.7	2055	4	AAK23453	AaK23453	10 gene n
716	13.8	76.7	1178	6	ABLA9811	AbL9811	Human pol	C	789	13.8	76.7	2055	4	AAI70258	Aai70258	Mouse TH
717	13.8	76.7	1185	5	AAH68383	Aah68383	C. glutami	C	790	13.8	76.7	2055	4	AAK52604	Aak52604	Murine TH
718	13.8	76.7	1251	4	AAK526446	Aak526446	Human CDN	C	791	13.8	76.7	2055	6	ABK53311	Abk53311	DNA encod
719	13.8	76.7	1251	4	ABX73787	Abx73787	Human nov	C	792	13.8	76.7	2055	9	ABQ77032	Abq77032	Murine T-
720	13.8	76.7	1255	4	AAK560857	Aak560857	Human can	C	793	13.8	76.7	2055	9	ADB37513	Adb37513	Mouse Th-
721	13.8	76.7	1276	3	AAC54737	Aac54737	Arabidops	C	794	13.8	76.7	2055	11	ADM68617	Adm68617	Different
722	13.8	76.7	1277	3	AAC54314	Aac54314	Arabidops	C	795	13.8	76.7	2091	9	ADB06513	AdB06513	Alloioioc
723	13.8	76.7	1299	5	ABV28991	Abv28991	Human pro	C	796	13.8	76.7	2091	9	ADB06509	AdB06509	Alloioioc
724	13.8	76.7	1299	5	ABV23150	Abv23150	Human pro	C	797	13.8	76.7	2091	4	ABO66511	AbO66511	Alloioioc
725	13.8	76.7	1299	6	ABZ14067	Abz14067	Arabidops	C	798	13.8	76.7	2158	4	AAI989108	Aai989108	Human neu
726	13.8	76.7	1299	12	ADN72146	Adn72146	Thale cre	C	799	13.8	76.7	2312	4	ABL24278	AbL24278	Drosophi
727	13.8	76.7	1302	11	ABD13749	Abd13749	Pseudomon	C	800	13.8	76.7	2312	8	ACA27258	Aca27258	Prokaryot
728	13.8	76.7	1325	3	AAC33679	Aac33679	Arabidops	C	801	13.8	76.7	2425	10	ADR79032	Adr79032	Human pro
729	13.8	76.7	1356	8	AAC51981	Aac51981	Arabidops	C	802	13.8	76.7	2437	4	AAK94582	Aak94582	Human ful
730	13.8	76.7	1356	8	ACA46498	Aca46498	Prokaryot	C	803	13.8	76.7	2437	12	ADL131475	AdL131475	Full leng
731	13.8	76.7	1359	4	AAH53220	Aah53220	S. epider	C	804	13.8	76.7	2439	6	ABO99310	AbO99310	Human cod
732	13.8	76.7	1392	6	ABN91627	Abn91627	Straphyloc	C	805	13.8	76.7	2540	4	ABL28220	AbL28220	Drosophi
733	13.8	76.7	1396	3	AAC39628	Aac39628	Arabidops	C	806	13.8	76.7	2553	9	ACH66191	Ach66191	Human tum
734	13.8	76.7	1399	4	ABL05733	AbL05733	Drosophi	C	807	13.8	76.7	2554	4	AAK27997	Aak27997	Human TRI
735	13.8	76.7	1399	4	AAK57083	Aak57083	CDNA enc	C	808	13.8	76.7	2554	6	AAD43683	Aad43683	Human TRI
736	13.8	76.7	1432	10	ADC35765	Adc35765	Drosophi	C	809	13.8	76.7	2610	5	AAK69775	Aak69775	Nucleotid
737	13.8	76.7	1435	3	AAC42309	Aac42309	Arabidops	C	810	13.8	76.7	2666	4	ABL26148	AbL26148	Drosophi
738	13.8	76.7	1435	3	AAC40033	Aac40033	Arabidops	C	811	13.8	76.7	2733	5	AAK69774	Aak69774	Nucleotid
739	13.8	76.7	1440	4	AAAD16713	Aaad16713	Cucumis m	C	812	13.8	76.7	2739	4	ABL09909	AbL09909	Drosophi
740	13.8	76.7	1440	4	AAC46559	Aac46559	Zea mays	C	813	13.8	76.7	2779	4	ABL07222	AbL07222	Drosophi
741	13.8	76.7	1470	6	ABZ16057	Abz16057	Arabidops	C	814	13.8	76.7	2781	2	AAQ13239	Aaq13239	Hsf CDNA
742	13.8	76.7	1502	4	AAD16683	Aad16683	Cucumis m	C	815	13.8	76.7	2781	2	AAQ25712	Aaq25712	Sequence
743	13.8	76.7	1576	3	AAC37497	Aac37497	Arabidops	C	816	13.8	76.7	2808	10	ADC90472	AdC90472	E. faecali
744	13.8	76.7	1586	4	AAK51766	Aak51766	Drosophi	C	817	13.8	76.7	2829	4	ABL26387	AbL26387	Drosophi
745	13.8	76.7	1589	12	ADI67069	Adi67069	Novel lac	C	818	13.8	76.7	2895	10	ADA54041	Ada54041	Human cod
746	13.8	76.7	1597	2	ABN50043	Abn50043	Enterococ	C	819	13.8	76.7	2918	4	ABL06526	AbL06526	Drosophi
747	13.8	76.7	1597	6	ABN58028	Abn58028	E. faecali	C	820	13.8	76.7	2936	12	ADK70733	AdK70733	Drosophi
748	13.8	76.7	1597	8	ACA87992	Aca87992	E. faecali	C	821	13.8	76.7	3001	10	ADK55029	AdK55029	Plant gro
749	13.8	76.7	1597	8	ABK61598	Abk61598	Enterococ	C	822	13.8	76.7	3027	12	ADN72586	Adn72586	Thale cre
750	13.8	76.7	1599	8	ABT19272	Abt19272	Aspergill	C	823	13.8	76.7	3096	4	ABL10095	AbL10095	Drosophi
751	13.8	76.7	1599	8	ABT21092	Abt21092	Aspergill	C	824	13.8	76.7	3138	4	ABL08323	AbL08323	Drosophi

825	13.8	76.7	3222	8	ACA25562	Prokaryot	898	13.8	76.7	3501	9	ADA97277	AdA97277 Human PRO
826	13.8	76.7	3260	3	AAAG5442	Human CAS	899	13.8	76.7	3501	9	ADB27034	AdB27034 cDNA enco
827	13.8	76.7	3294	4	AAH54973	S. epider	900	13.8	76.7	3501	9	ADB21967	AdB21967 Novel hum
828	13.8	76.7	3331	5	AAF89765	Nucleotid	901	13.8	76.7	3501	9	ADA66658	AdA66658 Human PRO
829	13.8	76.7	3331	10	ADF09659	Human Pre	902	13.8	76.7	3501	9	ADB22519	AdB22519 Human PRO
830	13.8	76.7	3334	4	AAF28030	Human TRL	903	13.8	76.7	3501	9	ADB23292	AdB23292 Human PRO
831	13.8	76.7	3334	6	AAF43716	Human TRL	904	13.8	76.7	3501	9	ADA92014	AdA92014 Novel hum
832	13.8	76.7	3386	4	ACH65224	Human tun	905	13.8	76.7	3501	9	ADB15077	AdB15077 Human PRO
833	13.8	76.7	3397	10	AAH5177	S. epider	906	13.8	76.7	3501	9	ADB38329	AdB38329 Human PRO
834	13.8	76.7	3397	10	ADB50622	Human pit	907	13.8	76.7	3501	9	ADB37777	AdB37777 Novel hum
835	13.8	76.7	3429	2	AAQ85413	Human JAK	908	13.8	76.7	3501	10	ADB66249	AdB66249 Novel hum
836	13.8	76.7	3429	3	AAQ85413	Human JAK	909	13.8	76.7	3501	10	ADB89329	AdB89329 Human PRO
837	13.8	76.7	3429	4	AAQ85413	Human JAK	910	13.8	76.7	3501	10	ADB90061	AdB90061 Human PRO
838	13.8	76.7	3429	8	ACC57650	Human pro	911	13.8	76.7	3501	10	ADB39162	AdB39162 Novel hum
839	13.8	76.7	3471	12	ADM79386	Human lym	912	13.8	76.7	3501	10	ADB44785	AdB44785 Novel hum
840	13.8	76.7	3501	4	AAH21262	Human cDN	913	13.8	76.7	3501	10	ADB86392	AdB86392 Human PRO
841	13.8	76.7	3501	8	ACA03621	cDNA enco	914	13.8	76.7	3501	10	ADB76997	AdB76997 Novel hum
842	13.8	76.7	3501	8	ABX89159	DNA enco	915	13.8	76.7	3501	10	ADB34154	AdB34154 Human PRO
843	13.8	76.7	3501	8	ACD41813	Human sec	916	13.8	76.7	3501	10	ADB35258	AdB35258 Human PRO
844	13.8	76.7	3501	8	ACA04042	Human cDN	917	13.8	76.7	3501	10	ADB33602	AdB33602 Human PRO
845	13.8	76.7	3501	8	ADA45556	Novel hum	918	13.8	76.7	3501	10	ADB34706	AdB34706 Human PRO
846	13.8	76.7	3501	9	ADA75987	Human PRO	919	13.8	76.7	3501	10	ADB35810	AdB35810 Human PRO
847	13.8	76.7	3501	9	ADA18637	Human PRO	920	13.8	76.7	3501	10	ADB46205	AdB46205 Novel hum
848	13.8	76.7	3501	9	ADA61260	Homo sapi	921	13.8	76.7	3501	10	ADC50078	AdC50078 Novel hum
849	13.8	76.7	3501	9	ADB19045	Novel hum	922	13.8	76.7	3501	10	ADC71625	AdC71625 Novel hum
850	13.8	76.7	3501	9	ADB27586	cDNA enco	923	13.8	76.7	3501	10	ADC59604	AdC59604 Novel hum
851	13.8	76.7	3501	9	ADA86065	Novel hum	924	13.8	76.7	3501	10	ADC53611	AdC53611 Novel hum
852	13.8	76.7	3501	9	ADB15629	Human PRO	925	13.8	76.7	3501	10	ADC56965	AdC56965 Novel hum
853	13.8	76.7	3501	9	ADA47415	Human PRO	926	13.8	76.7	3501	10	ADC60156	AdC60156 Novel hum
854	13.8	76.7	3501	9	ADA67210	Human PRO	927	13.8	76.7	3501	10	ADC50631	AdC50631 Novel hum
855	13.8	76.7	3501	9	ADB30217	cDNA enco	928	13.8	76.7	3501	10	ADC65158	AdC65158 Human PRO
856	13.8	76.7	3501	9	ADA85513	Novel hum	929	13.8	76.7	3501	10	ADC54256	AdC54256 Novel hum
857	13.8	76.7	3501	9	ADA96725	Human PRO	930	13.8	76.7	3501	10	ADC53217	AdC53217 Novel hum
858	13.8	76.7	3501	9	ADA79029	Human PRO	931	13.8	76.7	3501	10	ADC58740	AdC58740 Novel hum
859	13.8	76.7	3501	9	ADA87168	Novel hum	932	13.8	76.7	3501	10	ADC55618	AdC55618 Novel hum
860	13.8	76.7	3501	9	ADB16370	Human PRO	933	13.8	76.7	3501	10	ADC58188	AdC58188 Novel hum
861	13.8	76.7	3501	9	ADA91462	Novel hum	934	13.8	76.7	3501	10	ADC002862	AdC002862 Novel hum
862	13.8	76.7	3501	9	ADB14525	Human PRO	935	13.8	76.7	3501	10	ADC89854	AdC89854 Novel hum
863	13.8	76.7	3501	9	ADB18486	Novel hum	936	13.8	76.7	3501	10	ADC69273	AdC69273 cDNA enco
864	13.8	76.7	3501	9	ADA93701	Human PRO	937	13.8	76.7	3501	10	ADC48162	AdC48162 Human PRO
865	13.8	76.7	3501	9	ADB19597	Novel hum	938	13.8	76.7	3501	10	ADD09691	AdD09691 Human PRO
866	13.8	76.7	3501	9	ACD98442	Novel hum	939	13.8	76.7	3501	10	ADD04266	AdD04266 Novel hum
867	13.8	76.7	3501	9	ADA74163	Human PRO	940	13.8	76.7	3501	10	ADC80222	AdC80222 Novel hum
868	13.8	76.7	3501	9	ADB24396	Human PRO	941	13.8	76.7	3501	10	ADD10729	AdD10729 Human PRO
869	13.8	76.7	3501	9	ADA81920	Human PRO	942	13.8	76.7	3501	10	ADC47610	AdC47610 Human PRO
870	13.8	76.7	3501	9	ADA74883	Human PRO	943	13.8	76.7	3501	10	ADC79670	AdC79670 Novel hum
871	13.8	76.7	3501	9	ADA84961	Human PRO	944	13.8	76.7	3501	10	ADD09139	AdD09139 Human PRO
872	13.8	76.7	3501	9	ADA84409	Novel hum	945	13.8	76.7	3501	10	ADD040852	AdD040852 Novel hum
873	13.8	76.7	3501	9	ADB28665	cDNA enco	946	13.8	76.7	3501	10	ADD51991	AdD51991 cDNA enco
874	13.8	76.7	3501	9	ADA80193	Human PRO	947	13.8	76.7	3501	10	ADD54731	AdD54731 cDNA enco
875	13.8	76.7	3501	9	ADA75435	Human PRO	948	13.8	76.7	3501	10	ADD53283	AdD53283 Novel hum
876	13.8	76.7	3501	9	ADA46660	Human PRO	949	13.8	76.7	3501	10	ADD51439	AdD51439 cDNA enco
877	13.8	76.7	3501	9	ADA80745	Human PRO	950	13.8	76.7	3501	10	ADD02238	AdD02238 Human PRO
878	13.8	76.7	3501	9	ADA24956	Human PRO	951	13.8	76.7	3501	10	ADD01672	AdD01672 Human PRO
879	13.8	76.7	3501	9	ADA93132	Human PRO	952	13.8	76.7	3501	10	ADD53854	AdD53854 Novel hum
880	13.8	76.7	3501	9	ADB26482	cDNA enco	953	13.8	76.7	3501	10	ADD92171	AdD92171 Human PRO
881	13.8	76.7	3501	9	ADB30769	cDNA enco	954	13.8	76.7	3501	10	ADD91067	AdD91067 Human PRO
882	13.8	76.7	3501	9	ADA60697	Homo sapi	955	13.8	76.7	3501	10	ADE03681	AdE03681 Human PRO
883	13.8	76.7	3501	9	ADB23844	Human PRO	956	13.8	76.7	3501	10	ADE31978	AdE31978 Novel hum
884	13.8	76.7	3501	9	ADA96173	Human PRO	957	13.8	76.7	3501	10	ADE21910	AdE21910 cDNA enco
885	13.8	76.7	3501	9	ADA80745	Human PRO	958	13.8	76.7	3501	10	ADD79134	AdD79134 cDNA enco
886	13.8	76.7	3501	9	ADA95621	Human PRO	959	13.8	76.7	3501	10	ADE41670	AdE41670 Human PRO
887	13.8	76.7	3501	9	ADB25930	cDNA enco	960	13.8	76.7	3501	10	ADE17487	AdE17487 Human PRO
888	13.8	76.7	3501	9	ADB21415	Novel hum	961	13.8	76.7	3501	10	ADD91619	AdD91619 Human PRO
889	13.8	76.7	3501	9	ADA77194	Human PRO	962	13.8	76.7	3501	10	ADD33082	AdD33082 Novel hum
890	13.8	76.7	3501	9	ADB17934	cDNA enco	963	13.8	76.7	3501	10	ADE33634	AdE33634 Novel hum
891	13.8	76.7	3501	9	ADA86617	Novel hum	964	13.8	76.7	3501	10	ADD79686	AdD79686 cDNA enco
892	13.8	76.7	3501	9	ADA87720	Novel hum	965	13.8	76.7	3501	10	ADD92723	AdD92723 Human PRO
893	13.8	76.7	3501	9	ADA46108	Novel hum	966	13.8	76.7	3501	10	ADE19143	AdE19143 Human PRO
894	13.8	76.7	3501	9	ADB28138	cDNA enco	967	13.8	76.7	3501	10	ADE18591	AdE18591 Human PRO
895	13.8	76.7	3501	9	ADB28690	cDNA enco	968	13.8	76.7	3501	10	ADE42787	AdE42787 Human PRO
896	13.8	76.7	3501	9	ADA76642	Human PRO	969	13.8	76.7	3501	10	ADD95576	AdD95576 Human PRO
897	13.8	76.7	3501	9	ADA88272	Novel hum	970	13.8	76.7	3501	10	ADE22462	AdE22462 cDNA enco

971	13.8	76.7	3501	10	ADD85880	cDNA	enco
972	13.8	76.7	3501	10	AD352530	Novel	hum
973	13.8	76.7	3501	10	AD542222	Human	PRO
974	13.8	76.7	3501	10	ADD80238	cDNA	enco
975	13.8	76.7	3501	10	ADD89266	Human	PRO
976	13.8	76.7	3501	10	AD540550	Human	PRO
977	13.8	76.7	3501	10	AD5404349	Human	PRO
978	13.8	76.7	3501	10	AD592478	Human	PRO
979	13.8	76.7	3501	10	AD521187	Novel	hum
980	13.8	76.7	3501	10	AD522828	Novel	hum
981	13.8	76.7	3501	10	AD597163	Human	PRO
982	13.8	76.7	3501	10	AD580227	Human	PRO
983	13.8	76.7	3501	10	AD579675	Human	PRO
984	13.8	76.7	3501	10	AD545677	Novel	hum
985	13.8	76.7	3501	10	AD555519	Novel	hum
986	13.8	76.7	3501	10	AD567338	Novel	hum
987	13.8	76.7	3501	10	AD564687	Novel	hum
988	13.8	76.7	3501	10	AD563186	Novel	hum
989	13.8	76.7	3501	10	AD581600	Novel	hum
990	13.8	76.7	3501	10	AD581048	Novel	hum
991	13.8	76.7	3501	10	AD523871	Novel	hum
992	13.8	76.7	3501	10	AD567012	cDNA	enco
993	13.8	76.7	3501	11	AD582217	Novel	hum
994	13.8	76.7	3501	11	AD515616	Novel	hum
995	13.8	76.7	3501	11	AD516245	Novel	hum
996	13.8	76.7	3501	11	AD515664	Novel	hum
997	13.8	76.7	3501	11	AD514512	Novel	hum
998	13.8	76.7	3501	12	AD580774	Novel	hum
999	13.8	76.7	3501	12	AD576222	Human	PRO
1000	13.8	76.7	3501	12	AD587586	Human	PRO

ALIGNMENTS

Fragment Name	Begin	End
ABA90521_00	1	110000
ABA90521_01	100001	210000
ABA90521_02	200001	310000
ABA90521_03	300001	410000
ABA90521_04	400001	510000
ABA90521_05	500001	610000
ABA90521_06	600001	710000
ABA90521_07	700001	810000
ABA90521_08	800001	910000
ABA90521_09	900001	1010000
ABA90521_10	1000001	1110000
ABA90521_11	1100001	1210000
ABA90521_12	1200001	1310000
ABA90521_13	1300001	1410000
ABA90521_14	1400001	1510000
ABA90521_15	1500001	1610000
ABA90521_16	1600001	1710000
ABA90521_17	1700001	1810000
ABA90521_18	1800001	1910000
ABA90521_19	1900001	2010000
ABA90521_20	2000001	2110000
ABA90521_21	2100001	2210000
ABA90521_22	2200001	2310000
ABA90521_23	2300001	2365589

```

Query March          91.1%  Score 16.4; DB 6; Length 65589;
Best Local Similarity 94.4%  Pred. No. 4e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0
Qx      1 GAGAACACACAGTTCG 18
        |||||
Db      20471 GAGAACACACAGTTCG 20454

```

RESULT 2
 ID AEN84784
 AC AEN84784 standard; DNA; 19 BP.
 XX
 XX AEN84784;
 DT 05-NOV-2002 (first entry)
 XX
 XX Primer useful for familial dysautonomia allele genotype analysis.
 DE
 XX Familial dysautonomia; Riley-Day syndrome;
 XX hereditary sensory neuropathy III; human; carrier; diagnosis;
 XX IKAPAB kinase-complex associated protein; primer; SSCP;
 XX single-strand conformational polymorphism; ss.
 XX
 OS Homo sapiens.
 XX
 XX EPI252523-A2.
 FN
 XX 24-JUL-2002.
 PD
 XX
 XX 17-JAN-2002; 2002EP-00001232.
 PE
 XX
 XX 17-JAN-2001; 2001US-0262284P.
 PR
 XX
 XX (RUBI/) RUBIN B Y.
 PA (ANDE/) ANDERSON S L.
 XX
 XX Rubin BY, Anderson SL;
 PI
 XX WPI; 2002-601228/65.
 DR
 XX
 XX Detecting a polymorphism in a gene encoding the IkappaB kinase-complex-
 PT associated protein is used to diagnose and identify carriers of familial
 PT dysautonomia.
 XX
 XX Claim 7; Page 9; 16pp; English.
 PS
 XX
 XX The invention provides a method for detecting a polymorphism linked to a
 CC gene associated with familial dysautonomia (FD). This involves detecting
 CC a disruptive mutation in a gene encoding the IkappaB kinase-complex
 CC associated protein (IKAP) on chromosome 9q31. Sequence analysis of the
 CC IKAP-encoding gene showed, in chromosomes with the major FD haplotype, a
 CC T to C transition in position 6 of the donor splice site of intron 20.
 CC This mutation (2507+6T to C) results in the generation of an IKAP mRNA in
 CC which exon 20 is spliced out along with intron 20. Sequence analysis of
 CC the IKAP gene of individuals heterozygous for the FD chromosome with the
 CC most common minor haplotype (minor 2) showed a G to C transversion of
 CC nucleotide 2390 in exon 19 of the reported IKAP cDNA, resulting in an
 CC Arg96Ser amino acid substitution and disruption of a consensus Ser/Thr
 CC kinase phosphorylation site. The present sequence is a primer that can be
 CC used in a claimed method for detecting a disruptive mutation in the IKAP
 CC gene, using single-strand conformational polymorphism (SSCP) analysis.
 CC The primer was used in the genotype analysis of FD alleles. Anal with the
 CC primer given in AEN84785 yielded a 244 bp fragment. In a family with
 CC probands homozygous for the major haplotype, all affected individuals
 CC were homoallelic for 2507+6C to C and all the parents were heterozygous.
 CC In families with probands heterozygous for the major and minor 2 FD
 CC haplotypes, 1 parent and the proband were heterozygous for 2507+6T to C. Analysis
 CC of 31 probands homozygous for the major FD haplotype showed that 100% of
 CC the probands were homozygous for 2507+6T to C, 100% of the parents were
 CC heterozygous for this mutation, and 4 siblings of the probands had FD and
 CC were homozygous for the FD haplotype and the 2507+6T to C mutation. Study
 CC of a random group of 619 individuals of Ashkenazi Jewish descent revealed
 CC the presence of 25 carriers of 2507+6T to C and 2 individuals with R656P.
 CC The method is useful for FD diagnosis and for identifying carriers of the
 CC condition

Q	Sequence	19 BP;	8 A;	4 C;	4 G;	3 T;	0 U;	0 Other;
Query Match		88.9%;	Score	16;	DB	6;	Length	19;

Best Local Similarity 100.0%; Pred. No. 3.7e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAGAACACACAGATTTC 16
1 GAGAACACACAGATTTC 16
DB 1 GAGAACACACAGATTTC 16

RESULT 3

ABQ80567
ID ABQ80567 standard; DNA; 66479 BP.

AC ABQ80567;

DT 08-NOV-2002 (first entry)

DE Mutant human IKBKAP gene #2.

XX Human; IKBKAP; Familial Dysautonomia; FD; Riley-Day syndrome;

KW Hereditary Sensory and Autonomic Neuropathy Type III; carrier screening;

KM FD1; mutation; gene; chromosome 9q31; ds.

XX Homo sapiens.

OS Synthetic.

XX Key

FT mutation Location/Qualifiers
replace(33714,G)
/*tag= a

XX MO200259381-A2.

XX 01-AUG-2002.

XX 07-JAN-2002; 2002MO-US000473.

XX 06-JAN-2001; 2001US-026080P.

XX (GENO) GEN HOSPITAL CORP.

XX Slangenaupt S, Guseella JF;

XX WPI; 2002-674806/72.

XX Claim 1; Page; 109pp; English.

XX The present invention relates to methods and compositions useful for
CC detecting mutations which cause Familial Dysautonomia (FD, Riley-Day
CC syndrome, Hereditary Sensory and Autonomic Neuropathy Type III) (OMIM
CC 223900). It was found that mutations in the IKBKAP gene (see ABQ80565)
CC are associated with FD. The mutation associated with the major haplotype
CC of FD, FD1 mutation, is a base pair (bp) mutation, where the thymine
CC nucleotide located at bp 6 of intron 20 in the IKBKAP gene is replaced
CC with a cytosine. This results in skipping of exon 20 in the mRNA from FD
CC patients, although they continue to express varying levels of wild-type
CC message in a tissue-specific manner. The mutation associated with the
CC minor haplotype, FD2 mutation, is a bp mutation, where the guanine
CC nucleotide at bp 2397 (bp 73 of exon 19) is replaced with a cytosine.
CC This bp mutation causes an arginine to proline missense mutation (R696P)
CC in the IKBKAP protein, which is predicted to disrupt a potential
CC phosphorylation site. The IKBKAP nucleic acid sequences are useful for
CC identifying a subject with FD and for rapid carrier screening. The IKBKAP
CC gene contains 37 exons and maps to chromosome 9q31. Note: the present
CC sequence was not shown in the specification, but was derived from the
CC human wild-type IKBKAP sequence given in Fig 6

XX Sequence 66479 BP; 18271 A; 12399 C; 14128 G; 21681 T; 0 U; 0 Other;

Query Match 88.9%; Score 16; DB 6; Length 66479;

Best Local Similarity 100.0%; Pred. No. 6.3e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAGAACACACAGATTTC 16
1 GAGAACACACAGATTTC 33994
DB 33979 GAGAACACACAGATTTC 33994

RESULT 4

ABQ80566
ID ABQ80566 standard; DNA; 66479 BP.

AC ABQ80566;

DT 08-NOV-2002 (first entry)

DE Mutant human IKBKAP gene #1.

XX Human; IKBKAP; Familial Dysautonomia; FD; Riley-Day syndrome;

KW Hereditary Sensory and Autonomic Neuropathy Type III; carrier screening;

KM FD1; mutation; gene; chromosome 9q31; ds.

XX Homo sapiens.

OS Synthetic.

XX Key

FT mutation Location/Qualifiers
replace(34201,T)
/*tag= a

XX MO200259381-A2.

XX 01-AUG-2002.

XX 07-JAN-2002; 2002MO-US000473.

XX 06-JAN-2001; 2001US-026080P.

XX (GENO) GEN HOSPITAL CORP.

XX Slangenaupt S, Guseella JF;

XX WPI; 2002-674806/72.

XX Claim 1; Page; 109pp; English.

XX The present invention relates to methods and compositions useful for
CC detecting mutations which cause Familial Dysautonomia (FD, Riley-Day
CC syndrome, Hereditary Sensory and Autonomic Neuropathy Type III) (OMIM
CC 223900). It was found that mutations in the IKBKAP gene (see ABQ80565)
CC are associated with FD. The mutation associated with the major haplotype
CC of FD, FD1 mutation, is a base pair (bp) mutation, where the thymine
CC nucleotide located at bp 6 of intron 20 in the IKBKAP gene is replaced
CC with a cytosine. This results in skipping of exon 20 in the mRNA from FD
CC patients, although they continue to express varying levels of wild-type
CC message in a tissue-specific manner. The mutation associated with the
CC minor haplotype, FD2 mutation, is a bp mutation, where the guanine
CC nucleotide at bp 2397 (bp 73 of exon 19) is replaced with a cytosine.
CC This bp mutation causes an arginine to proline missense mutation (R696P)
CC in the IKBKAP protein, which is predicted to disrupt a potential
CC phosphorylation site. The IKBKAP nucleic acid sequences are useful for
CC identifying a subject with FD and for rapid carrier screening. The IKBKAP
CC gene contains 37 exons and maps to chromosome 9q31. Note: the present
CC sequence was not shown in the specification, but was derived from the
CC human wild-type IKBKAP sequence given in Fig 6

XX Sequence 66479 BP; 18271 A; 12399 C; 14129 G; 21680 T; 0 U; 0 Other;

Query Match 88.9%; Score 16; DB 6; Length 66479;

Best Local Similarity 100.0%; Pred. No. 6.3e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAGAACACACAGATTTC 16
|||||

DB 33979 GAGAACACACAGATTTC 33994

RESULT 5

ABQ80568 standard; DNA; 66479 BP.

XX AC ABO80568;

XX DT 08-NOV-2002 (first entry)

XX DE Mutant human IKKAP gene #3.

XX XX Human; IKKAP; Familial Dysautonomia; FD; Riley-Day syndrome;

XX KW Hereditary Sensory and Autonomic Neuropathy Type III; carrier screening;

XX KW FDI; mutation; gene; chromosome 9q31; ds.

XX OS Homo sapiens.

XX OS Synthetic.

XX FT mutation Location/Qualifiers

XX FT mutation replace(33714,G)

XX FT mutation /*tag= a

XX FT mutation replace(34201,T)

XX FT mutation /*tag= b

XX FT mutation

XX FT mutation

XX FT mutation

XX FT mutation

XX FT mutation

XX FT mutation

XX FT mutation

XX FT mutation

XX FT mutation

XX FT mutation

XX FT mutation

XX FT mutation

XX FT mutation

XX FT mutation

XX FT mutation

XX FT mutation

XX FT mutation

XX FT mutation

XX FT mutation

XX FT mutation

XX FT mutation

XX FT mutation

Query Match 88.9%; Score 16; DB 6; Length 66479;

Best Local Similarity 100.0%; Pred. No. 6.3e+02;

Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAGAACACACAGATTTC 16
|||||

DB 33979 GAGAACACACAGATTTC 33994

ABQ80565 standard; DNA; 66479 BP.

XX AC ABO80565;

XX DT 08-NOV-2002 (first entry)

XX DE Human IKKAP wild-type gene.

XX XX Human; IKKAP; Familial Dysautonomia; FD; Riley-Day syndrome;

XX KW Hereditary Sensory and Autonomic Neuropathy Type III; carrier screening;

XX KW gene; chromosome 9q31; ds.

XX OS Homo sapiens.

XX OS WO200259381-A2.

XX PD 01-AUG-2002.

XX PF 07-JAN-2002; 2002WO-US000473.

XX PR 06-JAN-2001; 2001US-026080P.

XX PA (GEO) GEN HOSPITAL CORP.

XX PI Slangenaupt S, Guseella JF;

XX DR WPI; 2002-674806/72.

XX PT New IKKAP genes with mutations, useful for identifying a subject with

XX PT Familial dysautonomia (FD), or for rapid carrier screening in the

XX PT Ashkenazi Jewish population, e.g. screening presymptomatic homozygotes or

XX PT prenatal diagnosis.

XX PS Claim 1; Fig 6; 109pp; English.

XX CC The present invention relates to methods and compositions useful for

XX CC detecting mutations which cause Familial Dysautonomia (FD, Riley-Day

XX CC syndrome, Hereditary Sensory and Autonomic Neuropathy Type III) (OMIM

XX CC 223900). It was found that mutations in the IKKAP gene (the present

XX CC sequence) are associated with FD. The mutation associated with the major

XX CC haplotype of FD, FDI mutation, is a base pair (bp) mutation, where the

XX CC thymine nucleotide located at bp 6 of intron 20 in the IKKAP gene is

XX CC replaced with a cytosine. This results in skipping of exon 20 in the mRNA

XX CC from FD patients, although they continue to express varying levels of

XX CC wild-type message in a tissue-specific manner. The mutation associated

XX CC with the minor haplotype, FD2 mutation, is a bp mutation, where the

XX CC guanine nucleotide at bp 2397 (bp 73 of exon 19) is replaced with a

XX CC mutation (R696P) in the IKKAP protein, which is predicted to disrupt a

XX CC potential phosphorylation site. The IKKAP nucleic acid sequences are

XX CC useful for identifying a subject with FD and for rapid carrier screening.

XX CC The IKKAP gene contains 37 exons and maps to chromosome 9q31

Sequence 66479 BP; 18271 A; 12398 C; 14129 G; 21681 T; 0 U; 0 Other;

Query Match 88.9%; Score 16; DB 6; Length 66479;

Best Local Similarity 100.0%; Pred. No. 6.3e+02;

Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAGAACACACAGATTTC 16
|||||

Db 33979 GAGAACACAGATTC 33994

RESULT 7

ABL37566 standard; cDNA, 485 BP.

ABL37566;

08-APR-2002 (first entry)

Human colon tumour antigen polynucleotide SEQ ID NO:1155.

Human; colon cancer; colon tumour antigen; cytosolic; vaccine;

colon tumour metastatic antigen; diagnosis; gene; ss.

Homo sapiens.

MO200196388-A2.

20-DEC-2001.

08-JUN-2001; 2001WO-US018557.

09-JUN-2000; 2000US-0210899P.

20-FEB-2001; 2001US-0270216P.

(CORI-) CORIXA CORP.

Jiang Y, Harlocker SL, Secrist H;

WPI; 2002-114514/15.

Novel isolated colon tumor polynucleotide differentially expressed in

colon tumor or colon metastatic tumor and polypeptides encoded by them,

useful for inhibiting development of cancer in patient.

Claim 1; SEQ ID NO 1155; 105bp; English.

AB136412 to AB138645 represent human colon tumour antigen cDNA clones (1)

which were isolated from human colon tumour and colon metastatic tumour

cDNA libraries. (1) have cytostatic activity and can be used in vaccine

production. (1) can be used for stimulating and/or expanding T cells

specific for a tumour protein on contact with the T cells. They are also

useful for inhibiting the development of cancer in a patient. (1) can be

used as probes or primers for nucleic acid hybridisation, for preparing

mutant species primers, or primers for use in genetic constructions. (1)

can be used in the diagnosis of a colon tumour

Sequence 485 BP; 121 A; 132 C; 126 G; 104 T; 0 U; 2 Other;

Query Match 85.6%; Score 15.4; DB 6; Length 485;

Best Local Similarity 94.1%; Pred. No. 9e+02; 1; Indels 0; Gaps 0;

Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

2 AGAACACACAGATTCGC 18

48 AGCACACACAGATTCGC 64

RESULT 8

AC32902

AC32902 standard; DNA, 513 BP.

AC32902;

19-JUN-2003 (first entry)

Prokaryotic essential gene #14559.

Antisense; ds; prokaryotic essential gene; cell proliferation;

drug design; gene.

xx

OS Enterococcus faecalis.

WO20027183-A2.

03-OCT-2002.

21-MAR-2002; 2002WO-US009107.

21-MAR-2001; 2001US-00815242.

06-SEP-2001; 2001US-00948993.

25-OCT-2001; 2001US-0342923P.

08-FEB-2002; 2002US-00072851.

06-MAR-2002; 2002US-0362699P.

(ELIT-) ELITRA PHARM INC.

Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen XL, Zyskind JW;

Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;

WPI; 2003-029926/02.

P-FSDB; AB029032.

New antisense nucleic acids, useful for identifying proteins or screening

for homologous nucleic acids required for cellular proliferation to

isolate candidate molecules for rational drug discovery programs.

Claim 14; SEQ ID NO 20772; 1766bp; English.

The invention relates to an isolated nucleic acid comprising any one of

the 6213 antisense sequences given in the specification where expression

of the nucleic acid inhibits proliferation of a cell. Also included are:

(1) a vector comprising a promoter operably linked to the nucleic acid

encoding a polypeptide whose expression is inhibited by the antisense

nucleic acid; (2) a host cell containing the vector; (3) an isolated

polypeptide or its fragment whose expression is inhibited by the

antisense nucleic acid; (4) an antibody capable of specifically binding

the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular

proliferation or the activity of a gene in an operon required for

proliferation; (7) identifying a compound that influences the activity of

the gene product or that has an activity against a biological pathway

required for proliferation, or that inhibits cellular proliferation; (8)

identifying a gene required for cellular proliferation or the biological

pathway in which a proliferation-required gene or its gene product lies

or a gene on which the test compound that inhibits proliferation of an

organism acts; (9) manufacturing an antibiotic; (10) profiling a

compound's activity; (11) a culture comprising strains in which the gene

product is overexpressed or underexpressed; (12) determining the extent

to which each of the strains is present in a culture or collection of

strains; or (13) identifying the target of a compound that inhibits the

proliferation of an organism. The antisense nucleic acids are useful for

identifying proteins or screening for homologous nucleic acids required

for cellular proliferation to isolate candidate molecules for rational

drug discovery programs, or for screening homologous nucleic acids

K. pneumoniae or P. aeruginosa. The present sequence is one of the target

prokaryotic essential genes. Note: The sequence data for this patent did

not form part of the printed specification, but was obtained in

electronic format directly from WIPO at

ftp.wipo.int/pub/published_pct_sequences

Sequence 513 BP; 171 A; 81 C; 115 G; 146 T; 0 U; 0 Other;

Query Match 85.6%; Score 15.4; DB 8; Length 513;

Best Local Similarity 94.1%; Pred. No. 9e+02; 1; Indels 0; Gaps 0;

Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

2 AGAACACACAGATTCGC 18

447 AGAACACACAGATTCGC 463

RESULT 9

ADH84638

ID ADH84638 standard; DNA, 534 BP.
XX
AC ADH84638;
XX
DT 22-APR-2004 (first entry)
XX
DE Enterococcus faecalis polynucleotide #2523.
XX
KW Enterococcus faecalis infection; transcription regulatory element;
KW antibacterial; gene; ds.
XX
OS Enterococcus faecalis.
XX
EN US6617156-B1.
XX
PD 09-SEP-2003.
XX
PF 13-AUG-1998; 98US-00134000.
XX
PR 15-AUG-1997; 97US-0055778P.
XX
PA (DOUC/) DOUCETTE-STAMM L A.
XX (BUSN/) BUSN D.
XX
PI Doucette-Stamm LA, Bush D;
XX
DR WP: 2003-895394/82.
DR P-PSDB; ADH86043.
XX
PT New nucleic acid comprising a sequence encoding an Enterococcus faecalis
PT polypeptide, useful for preparing a composition for diagnosing or
PT treating E. faecalis infection.
XX
PS Disclosure; SEQ ID NO 2523; 133bp; English.
XX
CC The invention relates to Enterococcus faecalis polynucleotides and
CC polypeptides. The invention also relates to a recombinant expression
CC vector comprising a polynucleotide operably linked to a transcription
CC regulatory element, a cell comprising a recombinant vector, a method for
CC producing an E. faecalis polypeptide, an isolated nucleic acid comprising
CC a sequence not given in the specification, a recombinant vector
CC comprising the nucleic acid and a cell comprising the recombinant vector.
CC The polynucleotides can be used to detect the presence of E. faecalis in
CC a sample. The sequences are useful for preparing a composition for
CC diagnosing or treating Enterococcus faecalis infection. This sequence
CC represents an E. faecalis polynucleotide of the invention.
XX
SQ Sequence 534 BP; 178 A; 82 C; 120 G; 154 T; 0 U; 0 Other;
Query Match 85.6%; Score 15.4; DB 10; Length 534;
Best Local Similarity 94.1%; Pred. No. 9e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 2 AGAACACACAGATTGCG 18
DB 465 AGAACACACAGATTGCG 481

RESULT 10
AAS52481/c
ID AAS52481 standard; DNA, 1470 BP.
XX
AC AAS52481;
XX
DT 13-FEB-2002 (first entry)
XX
DE E. coli DNA for cellular proliferation protein #203.
XX
KW Antisense; ds; prokaryotic cellular proliferation gene; antibiotic;
KW antibacterial; drug design.
XX
OS Escherichia coli.
XX

PN WO200170955-A2.
XX
PD 27-SEP-2001.
XX
PF 21-MAR-2001; 2001WO-US009180.
XX
PR 21-MAR-2000; 2000US-0191078P.
PR 23-MAY-2000; 2000US-0206848P.
PR 26-MAY-2000; 2000US-0207727P.
PR 23-OCT-2000; 2000US-024578P.
PR 27-NOV-2000; 2000US-0253625P.
PR 22-DEC-2000; 2000US-0257931P.
PR 16-FEB-2001; 2001US-0269308P.
XX
PA (ELIT-) ELITRA PHARM INC.
XX
PI Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
PI Yamamoto RT, Xu HH;
XX
DR WP: 2001-611495/70.
DR P-PSDB; AAU34622.
XX
PT New polynucleotides for the identification and development of
PT antibiotics, comprise sequences of antisense nucleic acids.
XX
PS Claim 27; SEQ ID NO 6118; 511bp; English.
XX
CC The invention relates to antisense inhibitors of genes essential to
CC prokaryotic cellular proliferation, their use in identifying the genes,
CC their use in the discovery of novel antibiotics, the essential genes
CC themselves and the encoded proteins. The prokaryotes used are *Escherichia*
CC *coli*, *Staphylococcus aureus*, *Salmonella typhi*, *Klebsiella pneumoniae*,
CC *Pseudomonas aeruginosa* and *Enterococcus faecalis*. The invention is also
CC useful for the identification of potential new targets for antibiotic
CC development. The antisense nucleic acids can also be used to identify
CC proteins used in proliferation, to express these proteins, and to obtain
CC antibodies capable of binding to the expressed proteins. The proteins can
CC be used to screen compounds in rational drug discovery programmes. The
CC antisense nucleic acid sequence is also useful to screen for homologous
CC nucleic acids which are required for cell proliferation in a wide variety
CC of organisms. The present sequence encodes an essential prokaryotic
CC cellular proliferation protein. Note: The sequence data for this patent
CC did not form part of the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pat_sequences
XX
SQ Sequence 1470 BP; 279 A; 366 C; 420 G; 405 T; 0 U; 0 Other;
Query Match 85.6%; Score 15.4; DB 4; Length 1470;
Best Local Similarity 94.1%; Pred. No. 9.6e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 2 AGAACACACAGATTGCG 18
DB 757 AGAACACACAGATTGCG 741

RESULT 11
ADE72695
ID ADE72695 standard; DNA, 1822 BP.
XX
AC ADE72695;
XX
DT 29-JAN-2004 (first entry)
XX
DE Human endometrial specific gene, SEQ ID NO 135.
XX
KW cytostatic; vaccine; human; endometrial specific genes;
KW endometrial specific protein; endometrial cancer; ds.
XX
OS Homo sapiens.
XX
PN WO200306081-A2.

XX PD 24-JUL-2003.
XX PF 23-DEC-2002; 2002MO-US041612.
XX PR 21-DEC-2001; 2001US-0342756P.
XX PA (DIAD-) DIADEXUS INC.
XX PI Sun Y, Liu C;
XX WPI; 2003-577666/54.
XX DR WPI; 2003-577666/54.
XX PT Nucleic acid molecules and polypeptides useful for diagnosing and
PT treating endometrial cancer and non-cancerous disease states in
PT endometrial.
XX PS Claim 1; SEQ ID NO 135; 824bp; English.
XX CC The invention comprises the amino acid and DNA sequences of human
CC endometrial specific genes and proteins. The DNA and protein sequences of
CC the invention are useful for diagnosing, imaging and treating a patient
CC with endometrial cancer. The present DNA sequence represents a human
CC endometrial specific gene of the invention.
XX SQ Sequence 1822 BP; 312 A; 578 C; 601 G; 331 T; 0 U; 0 Other;
Query Match 85.6%; Score 15.4; DB 10; Length 1822;
Best Local Similarity 94.1%; Pred. No. 9.8e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 2 AGAACACACAGATTGCG 18
Db 952 AGCACACACAGATTGCG 968
RESULT 12
AA88539/c
ID AA88539 standard; cDNA; 1944 BP.
XX AC AA88539;
XX DT 13-FEB-2002 (first entry)
XX DE DNA encoding novel human diagnostic protein #24343.
XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX OS Homo sapiens.
XX PN WO200175067-A2.
XX PD 11-OCT-2001.
XX PF 30-MAR-2001; 2001WO-US008631.
XX PR 31-MAR-2000; 2000US-00540217.
PR 23-AUG-2000; 2000US-00649167.
XX PA (HYSE-) HYSEQ INC.
XX PI Drmanac RT, Liu C, Tang YT;
XX DR WPI; 2001-639362/73.
DR P-PSDB; ABG24352.
XX PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX PS Claim 1; SEQ ID NO 24343; 103bp; English.

XX CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (II) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AA864197-AA894564 represent novel human diagnostic
CC coding sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX SQ Sequence 1944 BP; 419 A; 502 C; 506 G; 517 T; 0 U; 0 Other;
Query Match 85.6%; Score 15.4; DB 5; Length 1944;
Best Local Similarity 94.1%; Pred. No. 9.8e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 2 ABAACACACAGATTGCG 18
Db 307 AGAACACACAGATTGCG 291
RESULT 13
AA90079/c
ID AA90079 standard; cDNA; 1944 BP.
XX AC AA90079;
XX DT 13-FEB-2002 (first entry)
XX DE DNA encoding novel human diagnostic protein #25883.
XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX OS Homo sapiens.
XX PN WO200175067-A2.
XX PD 11-OCT-2001.
XX PF 30-MAR-2001; 2001WO-US008631.
XX PR 31-MAR-2000; 2000US-00540217.
PR 23-AUG-2000; 2000US-00649167.
XX PA (HYSE-) HYSEQ INC.
XX PI Drmanac RT, Liu C, Tang YT;
XX DR WPI; 2001-639362/73.
DR P-PSDB; ABG25892.
XX PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX PS Claim 1; SEQ ID NO 25883; 103bp; English.
XX CC The invention relates to isolated polynucleotide (I) and polypeptide (II)

CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
CC coding sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pcr_sequences

SQ Sequence 1944 BP, 419 A, 502 C, 506 G, 517 T, 0 U, 0 Other;
Query Match 85.6%; Score 15.4; DB 5; Length 1944;
Best Local Similarity 94.1%; Pred. No. 9.8e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 AGAACAAACAGATTGCG 18
DB 307 AGAACAAACAGATTGCG 291

RESULT 14
AAS93115/C
ID AAS93115 standard; cDNA; 1944 BP.
XX AAS93115;
XX
XX
XX 13-FEB-2002 (first entry)
XX
XX
XX DNA encoding novel human diagnostic protein #28919.
XX
XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
XX Homo sapiens.
XX
XX WO200175067-A2.
XX
XX
XX 11-OCT-2001.
XX
XX 30-MAR-2001; 2001WO-US008631.
XX
XX 31-MAR-2000; 2000US-00540217.
XX 23-AUG-2000; 2000US-00649167.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Drmanac RT, Liu C, Tang YT;
XX
XX WPI; 2001-639362/73.
XX P-PSDB; AAG28928.
XX
XX New isolated polynucleotide and encoded polypeptides, useful in
XX diagnostics, forensics, gene mapping, identification of mutations
XX responsible for genetic disorders or other traits and to assess
XX biodiversity.
XX
XX Claim 1; SEQ ID NO 28919; 103bp; English.
XX
XX The invention relates to isolated polynucleotide (I) and polypeptide (II)
XX sequences. (I) is useful as hybridisation probes, polymerase chain
XX reaction (PCR) primers, oligomers, and for chromosome and gene mapping,

CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
CC coding sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pcr_sequences

SQ Sequence 1944 BP, 419 A, 502 C, 506 G, 517 T, 0 U, 0 Other;
Query Match 85.6%; Score 15.4; DB 5; Length 1944;
Best Local Similarity 94.1%; Pred. No. 9.8e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 AGAACAAACAGATTGCG 18
DB 307 AGAACAAACAGATTGCG 291

RESULT 15
AAH48474/C
ID AAH48474 standard; DNA; 2186 BP.
XX AAH48474;
XX
XX
XX 21-SEP-2001 (first entry)
XX
XX
XX Escherichia coli lysP gene.
XX
XX
XX L-pipecolic acid production; delta-1-piperidine-6-carboxylic acid;
XX pyrrolidine-5-carboxylate reductase; lysine-6-aminotransferase; lysP, ds.
XX
XX Escherichia coli.
XX
XX WO200148216-A1.
XX
XX
XX 05-JUL-2001.
XX
XX 22-DEC-2000; 2000WO-JP009137.
XX
XX 28-DEC-1999; 99JP-00373389.
XX
XX (SAOC) MERCIAN CORP.
XX
XX Fujii T, Aritoku Y, Mukaihara M, Narita T, Agematsu H, Ishiki K;
XX
XX WPI; 2001-418288/44.
XX P-PSDB; AAG64107.
XX
XX Producing L-pipecolic acid for use as a pharmaceutical raw material,
XX comprises reducing delta-1-piperidine-6-carboxylic acid with pyrrolidine-
XX 5-carboxylate reductase particularly in recombinant bacterium.
XX
XX Disclosure; Page 35-39; 54pp; Japanese.
XX
XX The invention relates to a method for producing L-pipecolic acid. The
XX method comprises reducing delta-1-piperidine-6-carboxylic acid with
XX pyrrolidine-5-carboxylate reductase. A recombinant Escherichia coli or
XX Corynebacterium bacterium that can express a gene encoding lysine-6-
XX aminotransferase is cultured in an L-lysine-containing medium in order to
XX produce L-pipecolic acid. L-pipecolic acid may be used as a
XX pharmaceutical raw material. The process is efficient, and since it is

CC conducted enzymatically, chirality of the final product can be retained.
CC The present sequence is the Escherichia coli lypE gene
XX
SQ Sequence 2186 BP; 444 A; 531 C; 598 G; 613 T; 0 U; 0 Other;
Query Match 85.6%; Score 15.4; DB 4; Length 2186;
Best Local Similarity 94.1%; Pred. No. 9.9e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
OY 2 AGAACACACAGATTCCG 18
DB 1281 AGAACACACAGATTCCG 1265
RESULT 16
ADE72674
ID ADE72674 standard; DNA; 2620 BP.
AC ADE72674;
XX
XX 29-JAN-2004 (first entry)
DT
XX
XX Human endometrial specific gene, SEQ ID NO 114.
DE
XX cytostratic; vaccine; human; endometrial specific genes;
KW endometrial specific protein; endometrial cancer; ds.
XX
XX Homo sapiens.
OS
XX WO2003060081-A2.
PN
XX 24-JUL-2003.
PD
XX 23-DEC-2002; 2002WO-US041612.
PF
XX 21-DEC-2001; 2001US-0342756P.
PR
XX (DIAD-) DIADEXUS INC.
PA
XX Sun Y, Liu C;
PI
XX WPI; 2003-577666/54.
DR
XX Nucleic acid molecules and polypeptides useful for diagnosing and
PT treating endometrial cancer and non-cancerous disease states in
PT endometrial.
XX
XX Claim 1; SEQ ID NO 114; 824pp; English.
PS
XX The invention comprises the amino acid and DNA sequences of human
CC endometrial specific genes and proteins. The DNA and protein sequences of
CC the invention are useful for diagnosing, imaging and treating a patient
CC with endometrial cancer. The present DNA sequence represents a human
CC endometrial specific gene of the invention.
DE
XX Sequence 2620 BP; 523 A; 803 C; 806 G; 488 T; 0 U; 0 Other;
SQ
Query Match 85.6%; Score 15.4; DB 10; Length 2620;
Best Local Similarity 94.1%; Pred. No. 1e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
OY 2 AGAACACACAGATTCCG 18
DB 952 AGAACACACAGATTCCG 968
RESULT 17
ADE72675
ID ADE72675 standard; DNA; 2747 BP.
AC ADE72675;
XX
XX 29-JAN-2004 (first entry)
DT

XX Human endometrial specific gene, SEQ ID NO 115.
DE
XX cytostratic; vaccine; human; endometrial specific genes;
KW endometrial specific protein; endometrial cancer; ds.
XX
XX Homo sapiens.
OS
XX WO2003060081-A2.
PN
XX 24-JUL-2003.
PD
XX 23-DEC-2002; 2002WO-US041612.
PF
XX 21-DEC-2001; 2001US-0342756P.
PR
XX (DIAD-) DIADEXUS INC.
PA
XX Sun Y, Liu C;
PI
XX WPI; 2003-577666/54.
DR
XX Nucleic acid molecules and polypeptides useful for diagnosing and
PT treating endometrial cancer and non-cancerous disease states in
PT endometrial.
XX
XX Claim 1; SEQ ID NO 115; 824pp; English.
PS
XX The invention comprises the amino acid and DNA sequences of human
CC endometrial specific genes and proteins. The DNA and protein sequences of
CC the invention are useful for diagnosing, imaging and treating a patient
CC with endometrial cancer. The present DNA sequence represents a human
CC endometrial specific gene of the invention.
DE
XX Sequence 2747 BP; 530 A; 795 C; 836 G; 586 T; 0 U; 0 Other;
SQ
Query Match 85.6%; Score 15.4; DB 10; Length 2747;
Best Local Similarity 94.1%; Pred. No. 1e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
OY 2 AGAACACACAGATTCCG 18
DB 952 AGAACACACAGATTCCG 968
RESULT 18
ADE72693
ID ADE72693 standard; DNA; 2789 BP.
AC ADE72693;
XX
XX 29-JAN-2004 (first entry)
DT
XX Human endometrial specific gene, SEQ ID NO 133.
DE
XX cytostratic; vaccine; human; endometrial specific genes;
KW endometrial specific protein; endometrial cancer; ds.
XX
XX Homo sapiens.
OS
XX WO2003060081-A2.
PN
XX 24-JUL-2003.
PD
XX 23-DEC-2002; 2002WO-US041612.
PF
XX 21-DEC-2001; 2001US-0342756P.
PR
XX (DIAD-) DIADEXUS INC.
PA
XX Sun Y, Liu C;
PI
XX WPI; 2003-577666/54.
DR

XX Nucleic acid molecules and polypeptides useful for diagnosing and
 PT treating endometrial cancer and non-cancerous disease states in
 PT endometrial.
 XX
 PS Claim 1; SEQ ID NO 133; 824bp; English.
 CC The invention comprises the amino acid and DNA sequences of human
 CC endometrial specific genes and proteins. The DNA and protein sequences of
 CC the invention are useful for diagnosing, imaging and treating a patient
 CC with endometrial cancer. The present DNA sequence represents a human
 CC endometrial specific gene of the invention.
 XX
 SQ Sequence 2789 BP; 560 A; 661 C; 839 G; 525 T; 0 U; 4 Other;
 QY
 Query Match 85.6%; Score 15.4; DB 10; Length 2789;
 Best Local Similarity 94.1%; Pred. No. 1e+03;
 Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 Db 2 AGAACAACAGATTGCG 18
 952 AGCAACAACAGATTGCG 968
 RESULT 19
 AAV30824
 ID AAV30824 standard; DNA; 3178 BP.
 AC AAV30824;
 XX
 DT 14-SEP-1998 (first entry)
 XX
 DE Mitochondrial glycerol-3-phosphate dehydrogenase Gnt2 gene.
 XX
 KM Glycerol-3-phosphate dehydrogenase; G3PDH; GUT2; Y1L155C; ss.
 XX
 OS Saccharomyces sp.
 XX
 FH Key Location/Qualifiers
 FT CDS 468..2312
 FT /*tag= a
 FT /EC_number= "1.1.99.5"
 XX
 XX W09821340-A1.
 XX
 PD 22-MAY-1998.
 XX
 PF 10-NOV-1997; 97MO-US020293.
 XX
 PR 13-NOV-1996; 96US-0030602P.
 XX
 PA (DUPO) DU PONT DE NEMOURS & CO E I.
 PA (GEMV) GENENCOR INT INC.
 XX
 PI Balthis BA, Gatenby AA, Haynie SL, Hsu AK, Lareau RD;
 XX
 DR WPI; 1998-297943/26.
 DR P-PSDB; AAW57326.
 XX
 PT Fermentative production of glycerol using recombinant host - containing
 PT genes for glycerol-3-phosphate dehydrogenase and/or glycerol-3-
 PT phosphatase.
 XX
 PS Disclosure; Page 26-27; 57pp; English.
 XX
 CC This gene, designated GUT2 or Y1L155C, codes for a mitochondrial glycerol
 CC -3-phosphate dehydrogenase (G3PDH, see AAW57327) that catalyses the
 CC conversion of dihydroxyacetone phosphate to glycerol-3-phosphate. The
 CC invention provides recombinant organisms that express G3PDH and/or
 CC glycerol-3-phosphatase (G3P) (see also AAW57324-32) useful for the
 CC production of glycerol from a variety of C-sources. A host cell is
 CC preferably transformed with a cassette containing a G3PDH gene and/or a
 CC G3P gene and then cultured in the presence of a mono-, oligo-,

CC polyaccharide or 1C-substrate. The glycerol obtained is used in
 CC cosmetics, liquid soaps, pharmaceuticals, lubricants and antifreezes; its
 CC esters are used in the oil and fat industries. The method produces
 CC glycerol rapidly and inexpensively without generation of polluting by-
 CC products
 XX
 SQ Sequence 3178 BP; 946 A; 710 C; 646 G; 876 T; 0 U; 0 Other;
 QY
 Query Match 85.6%; Score 15.4; DB 2; Length 3178;
 Best Local Similarity 94.1%; Pred. No. 1e+03;
 Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 Db 2 AGAACAACAGATTGCG 18
 2154 AGAACAACAGATTGCG 2170
 RESULT 20
 AAV42018
 ID AAV42018 standard; DNA; 3178 BP.
 AC AAV42018;
 XX
 DT 12-OCT-1998 (first entry)
 XX
 DE Glycerol-3-phosphate dehydrogenase gnt2 gene.
 XX
 KM Glycerol-3-phosphate dehydrogenase; gnt2 gene; 1,3-propanediol; ss.
 XX
 OS Saccharomyces sp.
 XX
 FH Key Location/Qualifiers
 FT CDS 468..2312
 FT /*tag= a
 XX
 XX W09821339-A1.
 XX
 PD 22-MAY-1998.
 XX
 PF 10-NOV-1997; 97MO-US020292.
 XX
 PR 13-NOV-1996; 96US-0030601P.
 XX
 PA (DUPO) DU PONT DE NEMOURS & CO E I.
 PA (GEMV) GENENCOR INT INC.
 XX
 PI Gatenby AA, Haynie SL, Nagarajan V, Nair RV, Nakamura CB;
 PI Payne MS, Picatagiro SK, Dias-Torres W, Hsu AK, Lareau RD;
 PI Trimbur DE, Whited GW;
 XX
 DR WPI; 1998-297942/26.
 DR P-PSDB; AAW30682.
 XX
 PT Fermentative production of 1,3-propanediol - by single organism
 PT containing cassette comprising specific genes, and capable of using
 PT inexpensive carbon sources.
 XX
 PS Claim 9; Page 50-52; 95pp; English.
 XX
 CC This DNA sequence includes the coding region of the gnt2 gene encoding a
 CC mitochondrial glycerol-3-phosphate dehydrogenase (see AAW30682) of
 CC Saccharomyces sp. This enzyme catalyses the conversion of
 CC dihydroxyacetone phosphate to glycerol-3-phosphate. A claimed method for
 CC production of 1,3-propanediol (1) comprises culturing a microorganism
 CC transformed with a cassette containing at least 1 of the genes (see
 CC AAV42012-21) for glycerol-3-phosphate dehydrogenase, glycerol-3-
 CC phosphatase, glycerol dehydratase and 1,3-propanediol oxidoreductase (see
 CC AAW30682-85). Any of these genes not present on the cassette must be
 CC present endogenously. The enzyme sequences may include substitutions,
 CC deletions and additions provided activity is not altered. A single
 CC recombinant organism can now be used for production of (1) from
 CC inexpensive C-sources (contrast use of glycerol or dihydroxyacetone)
 CC without causing pollution. (1) is a starting material for polyesters,

CC polyurethanes and cyclic compounds

XX Sequence 3178 BP; 946 A; 709 C; 647 G; 876 T; 0 U; 0 Other;

SQ Query Match 85.6%; Score 15.4; DB 2; Length 3178;

XX Best Local Similarity 94.1%; Pred. No. 1e+03; Mismatches 16; Conservative 0; Indels 1; Gaps 0;

QY 2 AGAACAACAAGATTCCG 18

DB 2154 AGAACAACAAGATTCCG 2170

RESULT 21

AAV35739 ID AAV35739 standard; DNA; 3178 BP.

XX AAV35739;

XX 28-SEP-1998 (first entry)

XX Klebsiella pneumoniae glycerol-3-phosphate dehydrogenase GUT2 gene.

XX GUT2 gene; glycerol-3-phosphate dehydrogenase; production;

XX 1,3-propanediol; recombinant; mitochondrial; ss.

XX Klebsiella pneumoniae.

XX Key Location/Qualifiers

XX CDS 468..2309

XX /tag= a /product= "glycerol-3-phosphate dehydrogenase"

XX MO9821341-A2.

XX 22-MAY-1998.

XX 13-NOV-1997; 97WO-US020873.

XX 13-NOV-1996; 96US-0030601P.

XX (GENV) GENENCOR INT INC.

XX Dunn-Coleman NS, Diaz-Torres M, Chase MW, Trimbur D;

XX WPI; 1998-297944/26.

XX P-PSDB; AAM60257.

XX New method for increasing production of 1,3-propanediol - comprises

XX fermentation of inexpensive carbon sources by microorganism expressing

XX dehydratase, used, e.g. to prolong half-life of enzyme.

XX Disclosure; Page 59-61; 133pp; English.

XX The sequence is that of the GUT2 gene which encodes mitochondrial

XX glycerol-3-phosphate dehydrogenase. It was used as part of a method of

XX fermentative production of 1,3-propanediol (1,3-pd), using an organism

XX comprising at least 1 gene encoding a dehydratase, is improved by

XX inserting into the host a gene encoding protein X and culturing the

XX transformant in presence of a carbon source (e.g. mono-, oligo- or poly-

XX saccharide or 1C substrate) convertible to 1,3-pd. 1,3-pd is a starting

XX material for polyesters, polyurethanes and cyclic compounds. 1,3-pd can

XX now be produced by a single recombinant organism from inexpensive carbon

XX sources such as glucose (rather than costly glycerol or

DB 2154 AGAACAACAAGATTCCG 2170

RESULT 22

AA80613 ID AA80613 standard; DNA; 3178 BP.

XX AA80613;

XX 29-SEP-1999 (first entry)

XX GUT2 gene.

XX GUT2, glycerol-3-phosphate dehydrogenase; FAD-dependent enzyme;

XX EC 1.1.99.5; glycerol; recombinant organism; transformation; YII155C;

XX glycerol biosynthetic pathway; expression cassette; 1-3 propanediol;

XX pharmaceutical compound; antifreeze solution; lubricant; polyurethane;

XX cyclic compound; fat and oil industry; polyester fiber; ss.

XX Saccharomyces sp.

XX Key Location/Qualifiers

XX CDS 468..2312

XX /tag= a /product= "glycerol-3-phosphate dehydrogenase"

XX /note= "mitochondrial"

XX MO9928480-A1.

XX 10-JUN-1999.

XX 02-DEC-1998; 98WO-US025551.

XX 02-DEC-1997; 97US-00982783.

XX (DUPC) DU PONT DE NEMOURS & CO E I.

XX (GENV) GENENCOR INT INC.

XX Nair RV, Payne MS, Trimbur DE, Valle F;

XX WPI; 1999-385384/32.

XX P-PSDB; AAY26168.

XX Recombinant organisms containing G3PDH and or G3P phosphatase.

XX Disclosure; Page 53-55; 84pp; English.

XX The present sequence is a GUT2 gene alternatively termed YII155C encoding

XX mitochondrial glycerol-3-phosphate dehydrogenase (G3PDH) which catalyses

XX the conversion of dihydroxyacetone phosphate to glycerol-3-phosphate.

XX This is used to produce glycerol from a recombinant organism by

XX transforming a suitable host cell with an expression cassette comprising

XX either one or both of the genes encoding G3PDH and G3P, where the host

XX cell has disruptions in either glycerol kinase or glycerol dehydrogenase

XX endogenous genes to prevent their active expression. The transformed host

XX cell is cultured with a carbon source and glycerol is recovered.

XX Compounds derived from the glycerol biosynthetic pathway like 1,3-

XX propanediol can also be produced. The method provides a rapid,

XX inexpensive and environment-friendly source of glycerol. Glycerol is used

XX in cosmetics, food, pharmaceuticals, lubricants, anti-freeze solutions,

XX fat and oil industry etc. 1,3-propanediol is used for the production of

XX polyester fibers and the manufacture of polyurethanes and cyclic

XX compounds

SQ Sequence 3178 BP; 946 A; 709 C; 647 G; 876 T; 0 U; 0 Other;

Query Match 85.6%; Score 15.4; DB 2; Length 3178;

Best Local Similarity 94.1%; Pred. No. 1e+03; Mismatches 16; Conservative 0; Indels 1; Gaps 0;

QY 2 AGAACAACAAGATTCCG 18

Db 2154 AGAGACAGATTCGC 2170

RESULT 23

ID ADH72187

ADH72187 standard; DNA; 3333 BP.

AC ADH72187;

XX 25-MAR-2004 (first entry)

DE Human gene of the invention NOV5.1a SEQ ID NO:1083.

XX ds; gene; human; cytosolic; immunomodulator; neuroprotective; nootropic;

KW anorectic; antidiabetic; antimicrobial; antiparasitic; gene therapy;

KW vaccine; cancer; cachexia; Alzheimer's disease; Parkinson's disease;

KW obesity; diabetes; infectious disease; metabolic syndrome X;

XX dyslipidaemia.

CS Homo sapiens.

XX WO2003102155-A2.

PD 11-DEC-2003.

XX 03-JUN-2003; 2003WO-US017430.

PF 03-JUN-2002; 2002US-0385120P.

XX 04-JUN-2002; 2002US-0385784P.

PR 05-JUN-2002; 2002US-0386041P.

PR 05-JUN-2002; 2002US-0386047P.

PR 06-JUN-2002; 2002US-0386376P.

PR 06-JUN-2002; 2002US-0386453P.

PR 06-JUN-2002; 2002US-0386843P.

PR 06-JUN-2002; 2002US-0387016P.

PR 07-JUN-2002; 2002US-0387096P.

PR 07-JUN-2002; 2002US-0386815P.

PR 07-JUN-2002; 2002US-0386932P.

PR 07-JUN-2002; 2002US-0386942P.

PR 07-JUN-2002; 2002US-0386971P.

PR 07-JUN-2002; 2002US-0387262P.

PR 08-JUN-2002; 2002US-0296960P.

PR 10-JUN-2002; 2002US-0387400P.

PR 10-JUN-2002; 2002US-0387535P.

PR 11-JUN-2002; 2002US-0387610P.

PR 11-JUN-2002; 2002US-0387625P.

PR 11-JUN-2002; 2002US-0387634P.

PR 11-JUN-2002; 2002US-0387668P.

PR 11-JUN-2002; 2002US-0387696P.

PR 11-JUN-2002; 2002US-0387702P.

PR 11-JUN-2002; 2002US-0387836P.

PR 11-JUN-2002; 2002US-0387859P.

PR 12-JUN-2002; 2002US-0387933P.

PR 12-JUN-2002; 2002US-0387960P.

PR 12-JUN-2002; 2002US-0387962P.

PR 12-JUN-2002; 2002US-0388096P.

PR 13-JUN-2002; 2002US-0388123P.

PR 14-JUN-2002; 2002US-0389118P.

PR 14-JUN-2002; 2002US-0389170P.

PR 14-JUN-2002; 2002US-0389144P.

PR 14-JUN-2002; 2002US-0389146P.

PR 17-JUN-2002; 2002US-0389729P.

PR 17-JUN-2002; 2002US-0389742P.

PR 18-JUN-2002; 2002US-0389844P.

PR 19-JUN-2002; 2002US-0390066P.

PR 19-JUN-2002; 2002US-0390209P.

PR 21-JUN-2002; 2002US-0390763P.

PR 17-JUL-2002; 2002US-0396706P.

PR 06-AUG-2002; 2002US-0401628P.

PR 09-AUG-2002; 2002US-0402156P.

PR 09-AUG-2002; 2002US-0402256P.

PR 09-AUG-2002; 2002US-0402389P.

PR 12-AUG-2002; 2002US-0402786P.

PR 12-AUG-2002; 2002US-0402816P.

PR 12-AUG-2002; 2002US-0402821P.

PR 12-AUG-2002; 2002US-0402832P.

PR 13-AUG-2002; 2002US-0403448P.

PR 13-AUG-2002; 2002US-0403459P.

PR 13-AUG-2002; 2002US-0403531P.

PR 13-AUG-2002; 2002US-0403532P.

PR 13-AUG-2002; 2002US-0403533P.

PR 13-AUG-2002; 2002US-0403536P.

PR 15-AUG-2002; 2002US-0406317P.

PR 15-AUG-2002; 2002US-0406317P.

PR 26-AUG-2002; 2002US-0406182P.

PR 26-AUG-2002; 2002US-0406355P.

PR 27-AUG-2002; 2002US-0406240P.

PR 12-SEP-2002; 2002US-0410084P.

PR 20-SEP-2002; 2002US-0412528P.

PR 23-SEP-2002; 2002US-0412731P.

PR 30-SEP-2002; 2002US-0414801P.

PR 30-SEP-2002; 2002US-0414839P.

PR 30-SEP-2002; 2002US-0414840P.

PR 30-SEP-2002; 2002US-0414949P.

PR 09-OCT-2002; 2002US-0417186P.

PR 09-OCT-2002; 2002US-0417406P.

PR 23-OCT-2002; 2002US-0420639P.

PR 28-OCT-2002; 2002US-0421156P.

PR 31-OCT-2002; 2002US-0422690P.

PR 01-NOV-2002; 2002US-0423130P.

PR 05-NOV-2002; 2002US-0423798P.

PR 05-NOV-2002; 2002US-0423798P.

PR 12-NOV-2002; 2002US-0425453P.

XX (CURA-) CURAGEN CORP.

XX Alsobrook JP, Alvarez E, Anderson DM, Boldog FI, Casman SJ;

PI Catterton E, Chappoval A, Crabtree-Bokor JR, Edinger SR, Ellerman K;

PI Ettenberg S, Gangoli EA, Gerlach VL, Gorman L, Gunther E, Guo X;

PI Gusev VI, Herrmann UL, Ji W, Kekuda R, Li L, Liu X, MacDougall JR;

PI MacIsachlan T, Malayanar UM, Mezick AJ, Millet I, Mishra VS;

PI Padigaru M, Patirajan M, Pena CE, Peyman JA, Raha D, Raetelli L;

PI Rieger DK, Rothenberg ME, Sciore P, Shenoy SG, Shinkets RA;

PI Smithson G, Spletek KA, Stone DJ, Vernet CM, Vose EZ, Zhong M;

PI Zhong H;

XX WPI, 2004-081935/08.

DR P-FSDB; ADH72188.

XX New NOVX polypeptides and nucleic acid molecules useful for preventing or

PT treating NOVX-associated disorders, e.g. cancer, diabetes, infection or

PT obesity, and in chromosome mapping, tissue typing or pharmacogenomics.

XX Example 51; SEQ ID NO 1083; 1880bp; English.

PS The invention relates to a novel isolated polypeptide (NOVX). A

XX polypeptide of the invention has cytosolic, immunomodulator, and

CC neuroprotective; nootropic, anorectic, antidiabetic, antimicrobial, and

CC antiparasitic activity, and may have a use in gene therapy, and as a

CC vaccine. The polypeptides are encoded by NOVX polynucleotides comprising

CC any of the 303 fully defined nucleotide sequences given in the

CC specification. The polypeptide is useful in the manufacture of a

CC medicament for treating a syndrome associated with a human disease. The

CC polypeptide, polynucleotide and antibody are useful in diagnosing.

CC treating or preventing NOVX-associated disorders, e.g. cancer, cachexia,

CC Alzheimer's disease, Parkinson's disease, obesity, diabetes, infectious

CC diseases, metabolic syndrome X or dyslipidaemias. The nucleic acids are

CC further used as hybridisation probes, in chromosome mapping, tissue

CC typing, preventive medicine, and pharmacogenomics. The present sequence

CC encodes a NOVX polypeptide of the invention.

XX

XX Sequence 3333 BP; 771 A; 977 C; 936 G; 649 T; 0 U; 0 Other;

Query Match 85.6%; Score 15.4; DB 12; Length 3333;

Best Local Similarity 94.1%; Pred. No. 1e+03;

Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 AGAACAACAGATTCCG 18
DB 307 AGCAACAAGATTCCG 323

RESULT 24

ADH72193
ID ADH72193 standard; DNA, 3333 BP.

AC ADH72193;

DT 25-MAR-2004 (first entry)

XX Human gene of the invention NOV51d SEQ ID NO:1089.

XX de; gene; human; cytosstatic; immunomodulator; neuroprotective; nootropic;

XX anorectic; antidiabetic; antimicrobial; antilipemic; gene therapy;

XX vaccine; cancer; cachexia; Alzheimer's disease; Parkinson's disease;

XX obesity; diabetes; infectious disease; metabolic syndrome X;

XX dyslipidaemia.

XX Homo sapiens.

XX WC2003102155-A2.

XX 11-DEC-2003.

PF 03-JUN-2003; 2003WC-US017430.

XX 03-JUN-2002; 2002US-0385120P.

PR 04-JUN-2002; 2002US-0385784P.

PR 05-JUN-2002; 2002US-0386041P.

PR 06-JUN-2002; 2002US-0386376P.

PR 06-JUN-2002; 2002US-0386453P.

PR 06-JUN-2002; 2002US-0386864P.

PR 06-JUN-2002; 2002US-0387016P.

PR 07-JUN-2002; 2002US-0386796P.

PR 07-JUN-2002; 2002US-0386816P.

PR 07-JUN-2002; 2002US-0386933P.

PR 07-JUN-2002; 2002US-0386942P.

PR 07-JUN-2002; 2002US-0386971P.

PR 07-JUN-2002; 2002US-0387262P.

PR 08-JUN-2002; 2002US-0296960P.

PR 10-JUN-2002; 2002US-0387400P.

PR 10-JUN-2002; 2002US-0387535P.

PR 11-JUN-2002; 2002US-0387625P.

PR 11-JUN-2002; 2002US-0387633P.

PR 11-JUN-2002; 2002US-0387668P.

PR 11-JUN-2002; 2002US-0387696P.

PR 11-JUN-2002; 2002US-0387702P.

PR 11-JUN-2002; 2002US-0387836P.

PR 11-JUN-2002; 2002US-0387859P.

PR 12-JUN-2002; 2002US-0387933P.

PR 12-JUN-2002; 2002US-0387934P.

PR 12-JUN-2002; 2002US-0387960P.

PR 12-JUN-2002; 2002US-0388022P.

PR 12-JUN-2002; 2002US-0388096P.

PR 13-JUN-2002; 2002US-0388123P.

PR 14-JUN-2002; 2002US-0388181P.

PR 14-JUN-2002; 2002US-0388182P.

PR 14-JUN-2002; 2002US-0388144P.

PR 14-JUN-2002; 2002US-0389146P.

PR 17-JUN-2002; 2002US-0389729P.

PR 17-JUN-2002; 2002US-0389742P.

PR 18-JUN-2002; 2002US-0389884P.

PR 19-JUN-2002; 2002US-0390006P.

PR 19-JUN-2002; 2002US-0390009P.

PR 21-JUN-2002; 2002US-0390763P.

PR 17-JUL-2002; 2002US-0396706P.

PR 06-AUG-2002; 2002US-0401628P.

PR 09-AUG-2002; 2002US-0402156P.

PR 09-AUG-2002; 2002US-0402256P.

PR 09-AUG-2002; 2002US-0402389P.

PR 12-AUG-2002; 2002US-0402786P.

PR 12-AUG-2002; 2002US-0402815P.

PR 12-AUG-2002; 2002US-0402821P.

PR 12-AUG-2002; 2002US-0402832P.

PR 13-AUG-2002; 2002US-0403448P.

PR 13-AUG-2002; 2002US-0403459P.

PR 13-AUG-2002; 2002US-0403531P.

PR 13-AUG-2002; 2002US-0403532P.

PR 13-AUG-2002; 2002US-0403563P.

PR 13-AUG-2002; 2002US-0406117P.

PR 15-AUG-2002; 2002US-0406117P.

PR 26-AUG-2002; 2002US-0406182P.

PR 27-AUG-2002; 2002US-0406240P.

PR 12-SEP-2002; 2002US-0410084P.

PR 20-SEP-2002; 2002US-0412528P.

PR 23-SEP-2002; 2002US-0412731P.

PR 30-SEP-2002; 2002US-0414801P.

PR 30-SEP-2002; 2002US-0414839P.

PR 30-SEP-2002; 2002US-0414840P.

PR 30-SEP-2002; 2002US-0414954P.

PR 09-OCT-2002; 2002US-0417186P.

PR 09-OCT-2002; 2002US-0417406P.

PR 23-OCT-2002; 2002US-0420639P.

PR 28-OCT-2002; 2002US-0421156P.

PR 31-OCT-2002; 2002US-0422690P.

PR 01-NOV-2002; 2002US-0423130P.

PR 05-NOV-2002; 2002US-0423798P.

PR 12-NOV-2002; 2002US-0425453P.

(CURA-) CURAGEN CORP.

XX Alisbrook JP, Alvarez E, Anderson DM, Boldog FL, Casman SJ;
PI Catterton E, Chapoval A, Crabtree-Bokor JR, Edinger SR, Ellerman K;
PI Ettenberg S, Gangoli EA, Gerlach VL, Gorman L, Gunther E, Guo X;
PI Gusev V, Hermann JU, Ji W, Kekuda R, Li L, Liu X, MacDougall UB;
PI MacLachlan T, Malayanar UM, Mezick AV, Millec T, Mishra VS;
PI Padigar M, Patirajan M, Pena CE, Payman JA, Rana D, Rastelli L;
PI Rieger DK, Rothenberg ME, Sciore P, Shenoy SG, Shmukets RA;
PI Smithson G, Spytek KA, Stone DJ, Vernet CAM, Voss EZ, Zhong M;
PI Zhong H;

XX WPI; 2004-061935/08.

DR P-PSDB; ADH72194.

XX New NOVX polypeptides and nucleic acid molecules useful for preventing or
PT treating NOVX-associated disorders, e.g. cancer, diabetes, infection or
PT obesity, and in chromosome mapping, tissue typing or pharmacogenomics.

XX Example 51; SEQ ID NO 1089; 1880bp; English.

XX The invention relates to a novel isolated polypeptide (NOVX). A
CC polypeptide of the invention has cytosstatic, immunomodulator,
CC neuroprotective, nootropic, anorectic, antidiabetic, antimicrobial, and
CC antilipemic activity, and may have a use in gene therapy, and as a
CC vaccine. The polypeptides are encoded by NOVX polynucleotides comprising
CC any of the 303 fully defined nucleotide sequences given in the
CC specification. The polypeptide is useful in the manufacture of a
CC medicament for treating a syndrome associated with a human disease. The
CC polypeptide, polynucleotide and antibody are useful in diagnosing,
CC treating or preventing NOVX-associated disorders, e.g. cancer, cachexia,
CC Alzheimer's disease, Parkinson's disease, obesity, diabetes, infectious
CC diseases, metabolic syndrome X or dyslipidaemias. The nucleic acids are
CC further used as hybridisation probes, in chromosome mapping, tissue
CC typing, preventive medicine, and pharmacogenomics. The present sequence
CC encodes a NOVX polypeptide of the invention.

XX Sequence 3333 BP; 771 A; 978 C; 936 G; 648 T; 0 U; 0 Other;

Query Match 85.6%; Score 15.4; DB 12; Length 3333;
Best Local Similarity 94.1%; Pred. No. 1e+03; 1; Indels 0; Gaps 0;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 AGACACACAGATTGCG 18
DB 307 AGACACACAGATTGCG 323

RESULT 25
ADE72673
ID ADE72673 standard; DNA; 3368 BP.

AC ADE72673;

DT 29-JAN-2004 (first entry)

DE Human endometrial specific gene, SEQ ID NO 113.

KM cytosstatic; vaccine; human; endometrial specific genes;
endometrial specific protein; endometrial cancer; ds.

OS Homo sapiens.

PN WO2003060081-A2.

PD 24-JUL-2003.

PF 23-DEC-2002; 2002WO-US041612.

PR 21-DEC-2001; 2001US-0342756P.

PA (DIAD-) DIADEXUS INC.

PI Sun Y, Liu C;

DR WPI; 2003-577666/54.

PT Nucleic acid molecules and polypeptides useful for diagnosing and
treating endometrial cancer and non-cancerous disease states in
endometrial.

PS Claim 1; SEQ ID NO 113; 824bp; English.

CC The invention comprises the amino acid and DNA sequences of human
endometrial specific genes and proteins. The DNA and protein sequences of
the invention are useful for diagnosing, imaging and treating a patient
with endometrial cancer. The present DNA sequence represents a human
endometrial specific gene of the invention.

SEQ Sequence 3368 BP; 694 A; 1018 C; 1026 G; 630 T; 0 U; 0 Other;

Query Match 85.6%; Score 15.4; DB 10; Length 3368;

Best Local Similarity 94.1%; Pred. No. 1e+03; 1; Indels 0; Gaps 0;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 AGACACACAGATTGCG 18
DB 952 AGACACACAGATTGCG 968

RESULT 26
ADE72686
ID ADE72686 standard; DNA; 3470 BP.

AC ADE72686;

DT 29-JAN-2004 (first entry)

DE Human endometrial specific gene, SEQ ID NO 126.

KM cytosstatic; vaccine; human; endometrial specific genes;
endometrial specific protein; endometrial cancer; ds.

XX Homo sapiens.
OS
PN WO2003060081-A2.
PD 24-JUL-2003.
PF 23-DEC-2002; 2002WO-US041612.
PR 21-DEC-2001; 2001US-0342756P.
PA (DIAD-) DIADEXUS INC.
PI Sun Y, Liu C;
DR WPI; 2003-577666/54.

PT Nucleic acid molecules and polypeptides useful for diagnosing and
treating endometrial cancer and non-cancerous disease states in
endometrial.

PS Claim 1; SEQ ID NO 126; 824bp; English.

CC The invention comprises the amino acid and DNA sequences of human
endometrial specific genes and proteins. The DNA and protein sequences of
the invention are useful for diagnosing, imaging and treating a patient
with endometrial cancer. The present DNA sequence represents a human
endometrial specific gene of the invention.

SEQ Sequence 3470 BP; 792 A; 886 C; 992 G; 800 T; 0 U; 0 Other;

Query Match 85.6%; Score 15.4; DB 10; Length 3470;
Best Local Similarity 94.1%; Pred. No. 1e+03; 1; Indels 0; Gaps 0;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 AGACACACAGATTGCG 18
DB 952 AGACACACAGATTGCG 968

RESULT 27
ADE72691
ID ADE72691 standard; DNA; 3741 BP.

AC ADE72691;

DT 29-JAN-2004 (first entry)

DE Human endometrial specific gene, SEQ ID NO 131.

KM cytosstatic; vaccine; human; endometrial specific genes;
endometrial specific protein; endometrial cancer; ds.

OS Homo sapiens.

PN WO2003060081-A2.

PD 24-JUL-2003.

PF 23-DEC-2002; 2002WO-US041612.

PR 21-DEC-2001; 2001US-0342756P.

PA (DIAD-) DIADEXUS INC.

PI Sun Y, Liu C;

DR WPI; 2003-577666/54.

PT Nucleic acid molecules and polypeptides useful for diagnosing and
treating endometrial cancer and non-cancerous disease states in
endometrial.

PS Claim 1; SEQ ID NO 131; 824bp; English.

XX The invention comprises the amino acid and DNA sequences of human
CC endometrial specific genes and proteins. The DNA and protein sequences of
CC the invention are useful for diagnosing, imaging and treating a patient
CC with endometrial cancer. The present DNA sequence represents a human
CC endometrial specific gene of the invention.

XX Sequence 3741 BP; 778 A; 1121 C; 1125 G; 717 T; 0 U; 0 Other;

Query Match 85.6%; Score 15.4; DB 10; Length 3741;
Best Local Similarity 94.1%; Pred. No. 1e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 AGAACAACAAGATTGCG 18
DB 952 AGCACAACAAGATTGCG 968

RESULT 28

ADD69641
ID ADD69641 standard; cDNA; 3858 BP.

AC ADD69641;

DT 15-JAN-2004 (first entry)

DE Human REMAP cDNA - SEQ ID 70.

XX human; receptor and membrane-associated protein; REMAP; cytoprotective;
XX antiarteriosclerotic; anticonvulsant; neurotropic; neuroprotective;
XX cerebroprotective; anti-HIV; anti-allergic; anti-inflammatory;
XX thymostimetic; cell proliferative; cancer; atherosclerosis; neurological;
XX epilepsy; Huntington's disease; stroke; immune; inflammatory; AIDS;
XX allergy; developmental; hypothyroidism; Cushing's syndrome; infection;
XX 68; gene.

OS Homo sapiens.

XX MO2003048305-A2.

PD 12-JUN-2003.

XX 13-NOV-2002; 2002W0-US036759.

XX 13-NOV-2001; 2001US-0333097P.

XX 15-NOV-2001; 2001US-0335274P.

XX 14-DEC-2001; 2001US-0340542P.

XX 18-DEC-2001; 2001US-0342166P.

XX 11-JAN-2002; 2002US-0347589P.

XX 14-JAN-2002; 2002US-0348687P.

(INCY-) INCYTE GENOMICS INC.

XX Duggan BM, Yang J, Gietzen KJ, Lee SY, Tang YT, Azimzai Y;
PI Chawla NK, Warren BA, Barroso I, Becha SD, Yue H, Lehr-Mason PM;
PI Thangavelu K, Lee S, Emerling BM, Kaple AB, Khare R, Baughn MR;
PI Gandhi AR, Iran UK, Richardson TW, Marguis JP, Lal PG, Forsythe IJ;
PI Lee EA, Swarnakar A, Kallick DA, Griffin JA, Elliott VS, Gorvad AB;
PI Haralla ADA, Ison CH, Jin P, Jiang X, Jackson AA, Bhatia U;
PI Burrill JD, Blake JF, Ho A, Zheng W, Gao J;

XX WPI: 2003-513744/48.
DR P-PSDB; ADD69594.

PT New human receptors and membrane-associated proteins (REMAP), useful for
PT diagnosing, treating and preventing diseases or conditions associated
PT with the aberrant REMAP expression e.g. cancer, AIDS, atherosclerosis, or
PT infections.

XX Claim 5; SEQ ID NO 70; 298bp; English.

XX The invention relates to a novel isolated polypeptide comprising a human

CC receptor and membrane-associated protein (REMAP). The polypeptide of the
CC invention demonstrates cytoprotective, antiarteriosclerotic, anticonvulsant,
CC neurotropic, neuroprotective, cerebroprotective, anti-HIV, anti-allergic,
CC anti-inflammatory and thymostimetic activities and may be useful for
CC treating and diagnosing various disorders including those which are cell
CC proliferative such as cancer and atherosclerosis, neurological including
CC epilepsy, Huntington's disease and stroke, immune/inflammatory
CC and particularly AIDS and allergies and developmental such as hypothyroidism
CC and Cushing's syndrome, as well as infections. The current sequence is
CC that of the human REMAP cDNA of the invention.

XX Sequence 3858 BP; 972 A; 961 C; 989 G; 936 T; 0 U; 0 Other;

Query Match 85.6%; Score 15.4; DB 10; Length 3858;
Best Local Similarity 94.1%; Pred. No. 1e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 AGAACAACAAGATTGCG 18
DB 368 AGCACAACAAGATTGCG 384

RESULT 29

ADE72694
ID ADE72694 standard; DNA; 3886 BP.

AC ADE72694;

DT 29-JAN-2004 (first entry)

DE Human endometrial specific gene, SEQ ID NO 134.

XX cytoprotective; vaccine; human; endometrial specific genes;

XX endometrial specific protein; endometrial cancer; ds.

XX Homo sapiens.

XX MO2003060081-A2.

PD 24-JUL-2003.

XX 23-DEC-2002; 2002W0-US041612.

XX 21-DEC-2001; 2001US-0342756P.

XX (DIAD-) DIADEXUS INC.

XX Sun Y, Liu C;

XX WPI: 2003-577666/54.

PT Nucleic acid molecules and polypeptides useful for diagnosing and
PT treating endometrial cancer and non-cancerous disease states in
PT endometrial.

PS Claim 1; SEQ ID NO 134; 824bp; English.

XX The invention comprises the amino acid and DNA sequences of human
CC endometrial specific genes and proteins. The DNA and protein sequences of
CC the invention are useful for diagnosing, imaging and treating a patient
CC with endometrial cancer. The present DNA sequence represents a human
CC endometrial specific gene of the invention.

XX Sequence 3886 BP; 846 A; 1090 C; 1123 G; 827 T; 0 U; 0 Other;

Query Match 85.6%; Score 15.4; DB 10; Length 3886;
Best Local Similarity 94.1%; Pred. No. 1e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 AGAACAACAAGATTGCG 18
DB 952 AGCACAACAAGATTGCG 968

```
RESULT 30
PR ADE72687
XX ID ADE72687 standard; DNA; 4067 BP.
XX AC
XX ADE72687;
XX DT 29-JAN-2004 (first entry)
XX DE Human endometrial specific gene, SEQ ID NO 127.
XX KW cytoskeletal; vaccine; human; endometrial specific genes;
XX endometrial specific protein; endometrial cancer; ds.
XX OS Homo sapiens.
XX PN WO2003060081-A2.
XX PD 24-JUL-2003.
XX PF 23-DEC-2002; 2002WO-US041612.
XX PR 21-DEC-2001; 2001US-0342756P.
XX PA (DIAD-) DIADEXUS INC.
XX PI Sun Y, Liu C;
XX PI MPI; 2003-577666/54.
XX DR Nucleic acid molecules and polypeptides useful for diagnosing and
XX PT treating endometrial cancer and non-cancerous disease states in
XX PT endometrial.
XX PS Claim 1; SEQ ID NO 127; 824bp; English.
XX CC The invention comprises the amino acid and DNA sequences of human
XX CC endometrial specific genes and proteins. The DNA and protein sequences of
XX CC the invention are useful for diagnosing, imaging and treating a patient
XX CC with endometrial cancer. The present DNA sequence represents a human
XX CC endometrial specific gene of the invention.
XX SQ Sequence 4067 BP; 930 A; 1062 C; 1144 G; 931 T; 0 U; 0 Other;
OY
OY 2 AGACACACAGATTGCG 18
DB 952 AGCACAACAGATTGCG 968

Query Match 85.6%; Score 15.4; DB 10; Length 4067;
Best Local Similarity 94.1%; Pred. No. 1e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

RESULT 31
ADE72692
XX ID ADE72692 standard; DNA; 4331 BP.
XX AC
XX ADE72692;
XX DT 29-JAN-2004 (first entry)
XX DE Human endometrial specific gene, SEQ ID NO 132.
XX KW cytoskeletal; vaccine; human; endometrial specific genes;
XX endometrial specific protein; endometrial cancer; ds.
XX OS Homo sapiens.
XX PN WO2003060081-A2.
XX PD 24-JUL-2003.
XX PF 23-DEC-2002; 2002WO-US041612.
```

```
XX 21-DEC-2001; 2001US-0342756P.
XX (DIAD-) DIADEXUS INC.
XX Sun Y, Liu C;
XX MPI; 2003-577666/54.
XX Nucleic acid molecules and polypeptides useful for diagnosing and
XX PT treating endometrial cancer and non-cancerous disease states in
XX PT endometrial.
XX PS Claim 1; SEQ ID NO 132; 824bp; English.
XX CC The invention comprises the amino acid and DNA sequences of human
XX CC endometrial specific genes and proteins. The DNA and protein sequences of
XX CC the invention are useful for diagnosing, imaging and treating a patient
XX CC with endometrial cancer. The present DNA sequence represents a human
XX CC endometrial specific gene of the invention.
SQ Sequence 4331 BP; 901 A; 1273 C; 1276 G; 881 T; 0 U; 0 Other;
OY
OY 2 AGACACACAGATTGCG 18
DB 952 AGCACAACAGATTGCG 968

RESULT 32
ADE72671
XX ID ADE72671 standard; DNA; 4404 BP.
XX AC
XX ADE72671;
XX DT 29-JAN-2004 (first entry)
XX DE Human endometrial specific gene, SEQ ID NO 111.
XX KW cytoskeletal; vaccine; human; endometrial specific genes;
XX endometrial specific protein; endometrial cancer; ds.
XX OS Homo sapiens.
XX PN WO2003060081-A2.
XX PD 24-JUL-2003.
XX PF 23-DEC-2002; 2002WO-US041612.
XX PR 21-DEC-2001; 2001US-0342756P.
XX PA (DIAD-) DIADEXUS INC.
XX PI Sun Y, Liu C;
XX PI MPI; 2003-577666/54.
XX Nucleic acid molecules and polypeptides useful for diagnosing and
XX PT treating endometrial cancer and non-cancerous disease states in
XX PT endometrial.
XX PS Claim 1; SEQ ID NO 111; 824bp; English.
XX CC The invention comprises the amino acid and DNA sequences of human
XX CC endometrial specific genes and proteins. The DNA and protein sequences of
XX CC the invention are useful for diagnosing, imaging and treating a patient
XX CC with endometrial cancer. The present DNA sequence represents a human
XX CC endometrial specific gene of the invention.
```

SQ Sequence 4404 BP; 1009 A; 1166 C; 1228 G; 1001 T; 0 U; 0 Other;
 Query Match 85.6%; Score 15.4; DB 10; Length 4404;
 Best Local Similarity 94.1%; Pred. No. 1e+03;
 Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 2 AGAACACACGATTGCG 18
 DB 952 AGCACACACGATTGCG 968
 RESULT 33
 ID ABR99971 strand; DNA; 4717 BP.
 AC ABR99971;
 XX
 XX 21-OCT-2002 (first entry)
 DT
 DE DNA encoding human secreted protein SSCP-54.
 KW Secreted protein; SSCP; human; cell proliferative disorder; cancer;
 KW keratosis; arteriosclerosis; atherosclerosis; cirrhosis; hepatitis;
 KW psoriasis; autoimmune disorder; inflammatory disorder; AIDS; arthritis;
 KW acquired immunodeficiency syndrome; adult respiratory distress syndrome;
 KW Addison's disease; allergy; asthma; osteoporosis; autoimmune thyroiditis;
 KW Crohn's disease; dermatitis; diabetes; Graves' disease; haemodialysis;
 KW glomerulonephritis; scleroderma; systemic lupus erythematosus; uveitis;
 KW systemic sclerosis; ulcerative colitis; infection; trauma; Pick disease;
 KW cardiovascular disorder; congestive heart failure; angina; epilepsy;
 KW hypertensive heart disease; neurological disorder; Alzheimer's disease;
 KW Parkinson's disease; amyotrophic lateral sclerosis; stroke; dementia;
 KW Huntington's disease; multiple sclerosis; neuromuscular disorder;
 KW metabolic disorder; endocrine disorder; toxic myopathy; mental disorder;
 KW schizophrenia; disorder; developmental disorder; anaemia; epilepsy;
 KW hypothyroidism; glaucoma; sensorineural hearing loss; cataract;
 KW transgenic animal; gene; ds.
 XX
 XX Homo sapiens.
 OS
 XX
 XX MO200248337-A2.
 XX
 XX 20-JUN-2002.
 XX
 XX
 PF 12-DEC-2001; 2001WO-US048517.
 XX
 PR 13-DEC-2000; 2000US-0255639P.
 PR 21-DEC-2000; 2000US-0257852P.
 PR 05-JAN-2001; 2001US-0260105P.
 PR 18-JAN-2001; 2001US-0262932P.
 PR 18-JAN-2001; 2001US-0263096P.
 PR 19-JAN-2001; 2001US-0263090P.
 PR 02-FEB-2001; 2001US-0265926P.
 XX
 PA (INCY-) INCYTE GENOMICS INC.
 XX
 PI Griffin JA, Yao MG, Dugan BM, Yue H, Ding L, Lal PG, Lee BA;
 PI Ramkumar U, Thangaveelu K, Xu Y, Lee S, Tang YT, Nguyen DB;
 PI Warren BA, Honcheil CD, Gietzen KJ, Baughn MR, Gandhi AR, Arvizu C;
 PI Walla NK, Lu Y, Elliott VM, Lu DM, Hatalia AWA, Azimzai Y;
 PI Khan FA, Tran UK;
 XX
 DR WPI: 2002-583509/62.
 DR P-PSDB: ABG69674.
 XX
 PT Novel human secreted proteins and polynucleotides for diagnosing,
 PT preventing or treating cell proliferative, autoimmune/inflammatory,
 PT cardiovascular, neurological and developmental disorders.
 XX
 PS Claim 5; Page 23-234; 234pp; English.
 CC The invention describes an isolated polypeptide chosen from secreted
 CC proteins (I), SSCP I-54, (I) and the polynucleotide encoding it (II) are

CC useful for screening a compound for effectiveness as an agonist or
 CC antagonist of (I) or compound that alters expression of (II). (I), the
 CC identified agonist and antagonist are useful for treating a disease or
 CC condition associated with altered expression of functional SSCP in a patient.
 CC An antibody specific to (I) is useful for detecting the presence of (I),
 CC purifying (I) from a sample and for diagnosing a condition or disease
 CC associated with expression of SSCP in a subject or in a biological
 CC sample. (I) and (II) are modulators of (I) are useful for diagnosis,
 CC treatment and prevention of cell proliferative disorders (e.g. cancer,
 CC keratosis, arteriosclerosis, atherosclerosis, cirrhosis, hepatitis and
 CC psoriasis), autoimmune/inflammatory disorders (e.g. acquired
 CC immunodeficiency syndrome (AIDS), adult respiratory distress syndrome,
 CC Addison's disease, allergy, asthma, osteoporosis, autoimmune
 CC thyroiditis, Crohn's disease, dermatitis, diabetes, Graves' disease,
 CC glomerulonephritis, arthritis, scleroderma, systemic lupus erythematosus,
 CC systemic sclerosis, ulcerative colitis, haemodialysis, uveitis; viral,
 CC bacterial, fungal, parasitic, protozoal, helminthic infections and
 CC trauma), cardiovascular disorders (e.g. congestive heart failure, angina,
 CC hypertensive heart disease), neurological disorders (e.g. Alzheimer's and
 CC Pick disease, Parkinson disease, amyotrophic lateral sclerosis, epilepsy,
 CC stroke, Huntington's disease, multiple sclerosis, dementia, neuromuscular
 CC disorders, metabolic, endocrine and toxic myopathies, mental disorders,
 CC schizophrenia disorders, and developmental disorders (e.g. anaemia,
 CC epilepsy, hypothyroidism, glaucoma, sensorineural hearing loss and
 CC cataract). (II) is useful for creating transgenic animals to model human
 CC disease and to detect and quantify gene expression in biopsied tissues in
 CC which expression of SSCP is correlated with disease. This sequence
 CC encodes a human secreted protein (SCP)
 XX
 SQ Sequence 4717 BP; 1175 A; 1205 C; 1246 G; 1091 T; 0 U; 0 Other;
 Query Match 85.6%; Score 15.4; DB 6; Length 4717;
 Best Local Similarity 94.1%; Pred. No. 1e+03; 1; Indels 0; Gaps 0;
 Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 2 AGAACACACGATTGCG 18
 DB 368 AGCACACACGATTGCG 384
 RESULT 34
 ID ADE72672 standard; DNA; 4751 BP.
 AC ADE72672;
 XX
 XX 29-JAN-2004 (first entry)
 DT
 XX
 DE Human endometrial specific gene, SEQ ID NO 112.
 XX
 KW cytoskeletal; vaccine; human; endometrial specific genes;
 KW endometrial specific protein; endometrial cancer; ds.
 XX
 XX Homo sapiens.
 OS
 XX
 XX MO2003060081-A2.
 XX
 PD 24-JUL-2003.
 XX
 PF 23-DEC-2002; 2002WO-US041612.
 PR 21-DEC-2001; 2001US-0342756P.
 XX
 PA (DIAD-) DIADEXUS INC.
 XX
 PI Sun Y, Liu C;
 PI
 DR WPI: 2003-577666/54.
 XX
 PT Nucleic acid molecules and polypeptides useful for diagnosing and
 PT treating endometrial cancer and non-cancerous disease states in
 PT endometrial.

PS	Claim 1; SEQ ID NO 112; 824bp; English.
XX	
CC	The invention comprises the amino acid and DNA sequences of human
CC	endometrial specific genes and proteins. The DNA and protein sequences of
CC	the invention are useful for diagnosing, imaging and treating a patient
CC	with endometrial cancer. The present DNA sequence represents a human
CC	endometrial specific gene of the invention.
XX	
SQ	Sequence 4751 BP; 1050 A; 1346 C; 1381 G; 974 T; 0 U; 0 Other;
Query Match	85.6%; Score 15.4; DB 10; Length 4751;
Best Local Similarity	94.1%; Pred No. 1e+03; Mismatches 0; Indels 0; Gaps 0;
Matches	16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Dc	2 AGACCAACAAGATTGCG 18 952 AGCACACAAGATTGCG 968
YY	
RESULT 35	
ID	ABK99957
XX	ABK99957 standard; DNA; 4759 BP.
AC	
XX	ABK99957;
DT	21-OCT-2002 (first entry)
XX	
DE	DNA encoding human secreted protein SCEP-40.
XX	
KM	Secreted protein; SCEP; human; cell proliferative disorder; cancer;
KM	keraetosis; arteriosclerosis; atherosclerosis; cirrhosis; hepatitis;
KM	psoriasis; autoimmune disorder; inflammatory disorder; AIDS; arthritis;
KM	acquired immunodeficiency syndrome; adult respiratory distress syndrome;
KM	Addison's disease; allergy; asthma; osteoporosis; autoimmune thyroiditis;
KM	Crohn's disease; dermatitis; diabetes; Graves' disease; haemodialysis;
KM	glomerulonephritis; scleroderma; systemic lupus erythematosus; uveitis;
KM	systemic sclerosis; ulcerative colitis; infection; trauma; Pick disease;
KM	cardiovascular disorder; congestive heart failure; angina; epilepsy;
KM	hypertensive heart disease; neurological disorder; Alzheimer's disease;
KM	Parkinson's disease; amyotrophic lateral sclerosis; stroke; dementia;
KM	Huntington's disease; multiple sclerosis; neuromuscular disorder;
KM	metabolic disorder; endocrine disorder; toxic myopathy; mental disorder;
KM	schizophrenic disorder; developmental disorder; anaemia; epilepsy;
KM	hypothyroidism; glaucoma; sensorineural hearing loss; cataract;
KM	transgenic animal; gene; ds.
XX	
OS	Homo sapiens.
XX	
WO	WO200248337-A2.
XX	
PR	20-JUN-2002.
XX	
PF	12-DEC-2001; 2001WO-US048517.
XX	
PR	13-DEC-2000; 2000US-0255639P.
PR	21-DEC-2000; 2000US-0257852P.
PR	05-JAN-2001; 2001US-0260105P.
PR	18-JAN-2001; 2001US-0262932P.
PR	18-JAN-2001; 2001US-0263096P.
PR	19-JAN-2001; 2001US-0263090P.
PR	02-FEB-2001; 2001US-0265926P.
XX	
XX	(INCY-) INCYTE GENOMICS INC.
PI	Giffin JA, Yao MG, Duggan BM, Yue H, Ding L, Lai PG, Lee EA;
PI	Ramkumar J, Thangavelu K, Xu Y, Lee S, Tang YT, Nguyen DB, Arvizu C,
PI	Martin BA, Honcheill CD, Gietzen KT, Baughn KR, Gandhi AR, Arvizu C,
PI	Walia NK, Lu Y, Elliott VM, Lu DM, Hafalla AJA, Azimzai Y,
PI	Xian FA, Tran UK;
DR	WPI; 2002-583509/62.
DR	P-PADB; ABG69660.

Novel human secreted proteins and polynucleotides for diagnosing, preventing or treating cell proliferative, autoimmune/inflammatory, cardiovascular, neurological and developmental disorders.

Claim 5; Page 224-225; 234pp; English.

The invention describes an isolated polypeptide chosen from secreted proteins (I), SECP 1-54. (I) and the polynucleotide encoding it (II) are useful for screening a compound for effectiveness as an agonist or antagonist of (I) or compound that alters expression of (II). (I), the identified agonist and antagonist are useful for treating a disease or condition associated altered expression of functional SECP in a patient. An antibody specific to (I) is useful for detecting the presence of (I), purifying (I) from a sample and for diagnosing a condition or disease associated with expression of SECP in a subject or in a biological sample. (I) and (II) and modulators of (I) are useful for diagnosis, treatment and prevention of cell proliferative disorders (e.g. cancer, keratosis, arteriosclerosis, atherosclerosis, cirrhosis, hepatitis and psoriasis), autoimmune/inflammatory disorders (e.g. acquired immunodeficiency syndrome (AIDS), adult respiratory distress syndrome, Addison's disease, allergies, asthma, osteoporosis, autoimmune thyroiditis, Crohn's disease, dermatitis, diabetes, Graves' disease, CC glomerulonephritis, arthritis, scleroderma, systemic lupus erythematosus, CC systemic sclerosis, ulcerative colitis, haemodialysis, uveitis/ viral, bacterial, fungal, parasitic, protozoal, helminthic infections and CC trauma), cardiovascular disorders (e.g. congestive heart failure, angina, hypertensive heart disease), neurological disorders (e.g. Alzheimer's and stroke disease, Parkinson disease, amyotrophic lateral sclerosis, epilepsy, CC Pick, Huntington's disease, multiple sclerosis, dementia, neuromuscular disorders, metabolic, endocrine and toxic myopathies, mental disorders, CC schizophrenic disorders, and developmental disorders (e.g. anaemia, epilepsy, hypothyroidism, glaucoma, sensorineural hearing loss and CC cataract). (II) is useful for creating transgenic animals to model human disease and to detect and quantify gene expression in biopsied tissues in which expression of SECP is correlated with disease. This sequence encodes a human secreted protein (SCBP)

Sequence 4759 BP; 1176 A; 1226 C; 1257 G; 1100 T; 0 U; 0 Other;

Query Match 85.6%; Score 15.4; DB 6; Length 4759;
 Best Local Similarity 94.1%; Pred. No. 1e+03;
 Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0.

QY 2 AGAACACACAGATTGCG 18
 |||||
 DB 368 AGCACACACAGATTGCG 384

RESULT 36
 AAL57275
 ID AAL57275 standard; DNA; 4762 BP.
 XX AC
 XX AAL57275;
 DT 17-OCT-2003 (first entry)
 XX
 XX LIG-1 'human modifier of p53 pathway' DNA.
 XX
 XX Human modifier; p53 pathway modulating agent; HM; central integrator;
 KW DNA damage; hypoxia; nucleotide deprivation; oncogene activation;
 KW cytoskeletal; apoptotic; angiogenic; cell proliferation disorder; cancer;
 KW LIG-1; gene; ds; membrane glycoprotein.
 XX
 XX Homo sapiens.
 XX
 XX Key Location/Qualifiers
 FH 15..3296
 FT CDS
 FT /*tag= a
 FT /product= "LIG-1 'human modifier of p53 pathway' protein"
 XX
 XX W02003035833-A2.
 XX
 XX 01-MAY-2003.

XX 21-OCT-2002; 2002MO-US033542.
PF
XX 22-OCT-2001; 2001US-038733P.
PR
XX 15-FEB-2002; 2002US-0357600P.
PR
XX
XX (EXEL-) EXELIXIS INC.
PA
XX Belvin M, Francis-Lang H, Plozman GD, Funke RP, Li D, Friedman L;
PI
XX WPI; 2003-441259/41.
DR
XX P-FSDB; AAO23110.
PT
XX Identifying a candidate p53 pathway modulating agent useful for treating
PT cancer by contacting an assay system comprising a purified HM polypeptide
PT or nucleic acid with a test agent.
XX
XX
XX Example 2; Page 149-152; 205pp; English.
PS
XX The invention relates to a novel method which comprises identifying a
CC candidate p53 pathway-modulating agent. The human modifier (HM)
CC polypeptides and polynucleotides of the invention are identified with the
CC use of Drosophila melanogaster orthologues which are known to modify the
CC p53 pathway. The human p53 protein normally functions as a central
CC integrator of signals including DNA damage, hypoxia, nucleotide
CC deprivation and oncogene activation. The HM molecules of the invention
CC demonstrate cytostatic activity, whilst the method may be useful for
CC identifying a candidate p53 pathway-modulating agent and modulating the
CC p53 pathway in a cell, as well as during the diagnosis and treatment of
CC an apoptotic, angiogenic or cell proliferation disorder, such as cancer.
CC The current sequence is that of the Lig-I membrane glycoprotein HM DNA of
CC the invention
XX
SQ Sequence 4762 BP; 1168 A; 1229 C; 1264 G; 1101 T; 0 U; 0 Other;
Query Match 85.6%; Score 15.4; DB 8; Length 4762;
Best Local Similarity 94.1%; Pred. No. 1e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Oy 2 AGACACACAGATTCCG 18
Db 382 AGCACACACAGATTCCG 398
RESULT 37
ADH72191
ID ADH72191 standard; DNA; 4762 BP.
XX
AC ADH72191;
XX
DT 25-MAR-2004 (first entry)
XX
DE Human gene of the invention NC051c SEQ ID NO:1087.
XX
KW ds: gene; human; cytosolic; immunomodulator; neuroprotective; nocrotropic;
KW anorectic; antidiabetic; antimicrobial; antineoplastic; gene therapy;
KW vaccine; cancer; cachexia; Alzheimer's disease; Parkinson's disease;
KW obesity; diabetes; infectious disease; metabolic syndrome X;
KW dyslipidaemia.
XX
XX Homo sapiens.
OS
XX WO2003102155-A2.
XX
XX 11-DEC-2003.
XX
XX 03-JUN-2003; 2003MO-US017430.
XX
XX 03-JUN-2002; 2002US-0385120P.
PR
XX 04-JUN-2002; 2002US-0385784P.
PR
XX 05-JUN-2002; 2002US-0386041P.
PR
XX 06-JUN-2002; 2002US-0386376P.
PR

PR 06-JUN-2002; 2002US-0386453P.
PR 06-JUN-2002; 2002US-0386864P.
PR 06-JUN-2002; 2002US-0387016P.
PR 07-JUN-2002; 2002US-0386796P.
PR 07-JUN-2002; 2002US-0386816P.
PR 07-JUN-2002; 2002US-0386931P.
PR 07-JUN-2002; 2002US-0386944P.
PR 07-JUN-2002; 2002US-0386971P.
PR 07-JUN-2002; 2002US-0387262P.
PR 08-JUN-2002; 2002US-0296960P.
PR 10-JUN-2002; 2002US-0387400P.
PR 10-JUN-2002; 2002US-0387535P.
PR 11-JUN-2002; 2002US-0387610P.
PR 11-JUN-2002; 2002US-0387625P.
PR 11-JUN-2002; 2002US-0387634P.
PR 11-JUN-2002; 2002US-0387668P.
PR 11-JUN-2002; 2002US-0387696P.
PR 11-JUN-2002; 2002US-0387702P.
PR 11-JUN-2002; 2002US-0387836P.
PR 11-JUN-2002; 2002US-0387859P.
PR 12-JUN-2002; 2002US-0387933P.
PR 12-JUN-2002; 2002US-0387934P.
PR 12-JUN-2002; 2002US-0388023P.
PR 12-JUN-2002; 2002US-0388095P.
PR 13-JUN-2002; 2002US-0388123P.
PR 13-JUN-2002; 2002US-0388118P.
PR 14-JUN-2002; 2002US-03891120P.
PR 14-JUN-2002; 2002US-03891144P.
PR 14-JUN-2002; 2002US-0389146P.
PR 17-JUN-2002; 2002US-0389729P.
PR 17-JUN-2002; 2002US-0389742P.
PR 18-JUN-2002; 2002US-0389844P.
PR 19-JUN-2002; 2002US-0390069P.
PR 19-JUN-2002; 2002US-0390209P.
PR 21-JUN-2002; 2002US-0390763P.
PR 17-JUL-2002; 2002US-0396706P.
PR 06-AUG-2002; 2002US-0401628P.
PR 09-AUG-2002; 2002US-0402156P.
PR 09-AUG-2002; 2002US-0402566P.
PR 09-AUG-2002; 2002US-0402786P.
PR 12-AUG-2002; 2002US-0402816P.
PR 12-AUG-2002; 2002US-0402821P.
PR 12-AUG-2002; 2002US-0402832P.
PR 13-AUG-2002; 2002US-0403448P.
PR 13-AUG-2002; 2002US-0403459P.
PR 13-AUG-2002; 2002US-0403531P.
PR 13-AUG-2002; 2002US-0403532P.
PR 13-AUG-2002; 2002US-0403563P.
PR 13-AUG-2002; 2002US-0406317P.
PR 15-AUG-2002; 2002US-0406317P.
PR 26-AUG-2002; 2002US-0406182P.
PR 26-AUG-2002; 2002US-0406355P.
PR 27-AUG-2002; 2002US-0406240P.
PR 12-SEP-2002; 2002US-0410084P.
PR 20-SEP-2002; 2002US-0412528P.
PR 23-SEP-2002; 2002US-0412731P.
PR 30-SEP-2002; 2002US-0414801P.
PR 30-SEP-2002; 2002US-0414839P.
PR 30-SEP-2002; 2002US-0414840P.
PR 30-SEP-2002; 2002US-0414954P.
PR 09-OCT-2002; 2002US-0417186P.
PR 09-OCT-2002; 2002US-0417406P.
PR 23-OCT-2002; 2002US-0420639P.
PR 28-OCT-2002; 2002US-0421156P.
PR 31-OCT-2002; 2002US-0422690P.
PR 01-NOV-2002; 2002US-0423130P.
PR 05-NOV-2002; 2002US-0423798P.
PR 12-NOV-2002; 2002US-0425453P.
XX
XX (CURA-) CURAGEN CORP.
PA

Mon Dec 6 12:24:34 2004

us-10-050-189a-6.rng

Page 27

[illegible]

(CURA-) CURAGEN CORP

PI Rastellil L, Mezes PD, Smithson G, Guo X, Gerlach V, Casman SJ; CAM
PI Boldt-Fill L, Li L, Zernhusen BP, Tichener VT, Gangoli EA, Vernet
PI Pena CEA, Bulgures C, Liu X, Spyker KA, Gorman L, Spaderna SK,
PI Voss EA, Malyankar UM, Anderson DW, Pattanaian M, Miller CE;
PI Timpler RJ, Padigaru M, Shenoy SG, Kekuda R, Gusev VV, Pochart PF,
PI Zhong M,
DR W21, 2002-732824/79.
DR P-PSDB, ABB70116.

P-PSDB; ABP70116.

PT New NOV polypeptides and polynucleotides, useful for preventing, PT diagnosing or treating NOV-associated disorders e.g. diabetes, cancer PT Alzheimer's disease, dyslipidemias, obesity, immune or hematopoietic PT disorders, and asthma.

Claim 16; Page 208-210; 619pp; English.

The present invention relates to new isolated proteins (NOVX) and their NOVX coding sequences (ABV99327-ABV99595 and ABV70049-ABV70149), where X is any number from 1 to 48. The NOVX proteins and coding sequences are useful in the manufacture of a medicament for treating a syndrome associated with a human disease, preferably a NOVX-associated disorder. The NOVX coding sequences and proteins are useful for treating/preventing or diagnosing diseases such as metabolic disorders, diabetes, obesity, infectious disease, anorexia, cancer-associated cachexia, cancer, neurodegenerative diseases, Alzheimer's disease, Parkinson's disease, immunodeficiencies, haemopoietic disorders, cardiovascular disorders, fertility, bronchial asthma, AIDS, dyslipidaemia, metabolic disturbances associated with obesity, metabolic syndrome X or wasting disorders associated with chronic diseases or various cancers. The NOVX coding sequences and proteins may also be used as targets for the identification of small molecules that modulate or inhibit e.g. neurogenesis, cell differentiation, cell proliferation, haematopoiesis, wound healing and angiogenesis, in gene therapy, in generation of antibodies that bind immunospecifically to NOVX substances for use in therapeutic or diagnostic methods.

SQ Sequence 5101 BP; 1271 A; 1301 C; 1309 G; 1220 T; 0 U; 0 Other;

Query Match	85.6%	Score 15.4	DB 6	Length 5101
Best Local Similarity	94.1%	Pred. No. 1e+03		
Matches	16	Conservative	0	Mismatches 1; Indels 0; Gaps 0
QY	2	AGAACACACAGATTGCG	18	
Db	396	AGCACACACAGATTGCG	412	

396 AGCACAACAAGATTCCG 412

RESULT	39
ADH72189	
ID	ADH72189 standard; DNA; 5101 B2
XX	
AC	ADH72189;
XX	
DT	25-MAR-2004 (first entry)
XX	

DE Human gene of the invention NOV51b SEQ ID NO:1085

KW ds, gene; human; cytostatic; immunomodulator; neuroprotective; nootropic
KW anorectic; antidiabetic; antimicrobial; antihypaemic; gene therapy;
KW vaccine; cancer; cachexia; Alzheimer's disease; Parkinson's disease;
KW obesity; diabetes; infectious disease; metabolic syndrome X;
KW dyslipidaemia.

OS Homo sapiens.

PN WO2003102155-A2

PD 11-DEC-2003

PF 03-JUN-2003; 2003WO-US017430.
XX

PR 03-JUN-2002; 2002US-0385120P.

PR 05-JUN-2002; 2002US-0386041P;
DE 05-JUN-2002 0300UTC 0300047P

PR 06-JUN-2002; 2002US-0386376P.
06-JUN-2002; 2002US-0386453P.

PR 06-JUN-2002; 2002US-0386864P
DE 06-JUN-2002; 2002US-0387015P

PR 07-JUN-2002; 2002US-0386/96P.
PR 07-JUN-2002; 2002US-0386/96P.

PR 07-JUN-2002; 2002US-0386931P.
07-JUN-2002; 2002US-0386942P

PR 07-JUN-2002; 2002US-03869/1P.
PR 07-JUN-2002; 2002US-0387262P.

PR 10-JUN-2002; 2002HS-0387400P

11-JUN-2002: 2002US-0387610P
10-JUN-2002: 2002US-0387533P
09-JUN-2002: 2002US-0387456P
08-JUN-2002: 2002US-0387379P
07-JUN-2002: 2002US-0387302P
06-JUN-2002: 2002US-0387225P
05-JUN-2002: 2002US-0387148P
04-JUN-2002: 2002US-0387071P
03-JUN-2002: 2002US-0386994P
02-JUN-2002: 2002US-0386917P
01-JUN-2002: 2002US-0386840P
31-MAY-2002: 2002US-0386763P
30-MAY-2002: 2002US-0386686P
29-MAY-2002: 2002US-0386609P
28-MAY-2002: 2002US-0386532P
27-MAY-2002: 2002US-0386455P
26-MAY-2002: 2002US-0386378P
25-MAY-2002: 2002US-0386301P
24-MAY-2002: 2002US-0386224P
23-MAY-2002: 2002US-0386147P
22-MAY-2002: 2002US-0386070P
21-MAY-2002: 2002US-0385993P
20-MAY-2002: 2002US-0385916P
19-MAY-2002: 2002US-0385839P
18-MAY-2002: 2002US-0385762P
17-MAY-2002: 2002US-0385685P
16-MAY-2002: 2002US-0385608P
15-MAY-2002: 2002US-0385531P
14-MAY-2002: 2002US-0385454P
13-MAY-2002: 2002US-0385377P
12-MAY-2002: 2002US-0385300P
11-MAY-2002: 2002US-0385223P
10-MAY-2002: 2002US-0385146P
09-MAY-2002: 2002US-0385069P
08-MAY-2002: 2002US-0384992P
07-MAY-2002: 2002US-0384915P
06-MAY-2002: 2002US-0384838P
05-MAY-2002: 2002US-0384761P
04-MAY-2002: 2002US-0384684P
03-MAY-2002: 2002US-0384607P
02-MAY-2002: 2002US-0384530P
01-MAY-2002: 2002US-0384453P
30-APR-2002: 2002US-0384376P
29-APR-2002: 2002US-0384299P
28-APR-2002: 2002US-0384222P
27-APR-2002: 2002US-0384145P
26-APR-2002: 2002US-0384068P
25-APR-2002: 2002US-0383991P
24-APR-2002: 2002US-0383914P
23-APR-2002: 2002US-0383837P
22-APR-2002: 2002US-0383760P
21-APR-2002: 2002US-0383683P
20-APR-2002: 2002US-0383606P
19-APR-2002: 2002US-0383529P
18-APR-2002: 2002US-0383452P
17-APR-2002: 2002US-0383375P
16-APR-2002: 2002US-0383298P
15-APR-2002: 2002US-0383221P
14-APR-2002: 2002US-0383144P
13-APR-2002: 2002US-0383067P
12-APR-2002: 2002US-0382990P
11-APR-2002: 2002US-0382913P
10-APR-2002: 2002US-0382836P
09-APR-2002: 2002US-0382759P
08-APR-2002: 2002US-0382682P
07-APR-2002: 2002US-0382605P
06-APR-2002: 2002US-0382528P
05-APR-2002: 2002US-0382451P
04-APR-2002: 2002US-0382374P
03-APR-2002: 2002US-0382297P
02-APR-2002: 2002US-0382220P
01-APR-2002: 2002US-0382143P
31-MAR-2002: 2002US-0382066P
30-MAR-2002: 2002US-0381989P
29-MAR-2002: 2002US-0381912P
28-MAR-2002: 2002US-0381835P
27-MAR-2002: 2002US-0381758P
26-MAR-2002: 2002US-0381681P
25-MAR-2002: 2002US-0381604P
24-MAR-2002: 2002US-0381527P
23-MAR-2002: 2002US-0381450P
22-MAR-2002: 2002US-0381373P
21-MAR-2002: 2002US-0381296P
20-MAR-2002: 2002US-0381219P
19-MAR-2002: 2002US-0381142P
18-MAR-2002: 2002US-0381065P
17-MAR-2002: 2002US-0380988P
16-MAR-2002: 2002US-0380911P
15-MAR-2002: 2002US-0380834P
14-MAR-2002: 2002US-0380757P
13-MAR-2002: 2002US-0380680P
12-MAR-2002: 2002US-0380603P
11-MAR-2002: 2002US-0380526P
10-MAR-2002: 2002US-0380449P
09-MAR-2002: 2002US-0380372P
08-MAR-2002: 2002US-0380295P
07-MAR-2002: 2002US-0380218P
06-MAR-2002: 2002US-0380141P
05-MAR-2002: 2002US-0380064P
04-MAR-2002: 2002US-0379987P
03-MAR-2002: 2002US-0379910P
02-MAR-2002: 2002US-0379833P
01-MAR-2002: 2002US-0379756P
29-FEB-2002: 2002US-0379679P
28-FEB-2002: 2002US-0379602P
27-FEB-2002: 2002US-0379525P
26-FEB-2002: 2002US-0379448P
25-FEB-2002: 2002US-0379371P
24-FEB-2002: 2002US-0379294P
23-FEB-2002: 2002US-0379217P
22-FEB-2002: 2002US-0379140P
21-FEB-2002: 2002US-0379063P
20-FEB-2002: 2002US-0378986P
19-FEB-2002: 2002US-0378909P
18-FEB-2002: 2002US-0378832P
17-FEB-2002: 2002US-0378755P
16-FEB-2002: 2002US-0378678P
15-FEB-2002: 2002US-0378601P
14-FEB-2002: 2002US-0378524P
13-FEB-2002: 2002US-0378447P
12-FEB-2002: 2002US-0378370P
11-FEB-2002: 2002US-0378293P
10-FEB-2002: 2002US-0378216P
09-FEB-2002: 2002US-0378139P
08-FEB-2002: 2002US-0378062P
07-FEB-2002: 2002US-0377985P
06-FEB-2002: 2002US-0377908P
05-FEB-2002: 2002US-0377831P
04-FEB-2002: 2002US-0377754P
03-FEB-2002: 2002US-0377677P
02-FEB-2002: 2002US-0377600P
01-FEB-2002: 2002US-0377523P
31-JAN-2002: 2002US-0377446P
30-JAN-2002: 2002US-0377369P
29-JAN-2002: 2002US-0377292P
28-JAN-2002: 2002US-0377215P
27-JAN-2002: 2002US-0377138P
26-JAN-2002: 2002US-0377061P
25-JAN-2002: 2002US-0376984P
24-JAN-2002: 2002US-0376907P
23-JAN-2002: 2002US-0376830P
22-JAN-2002: 2002US-0376753P
21-JAN-2002: 2002US-0376676P
20-JAN-2002: 2002US-0376599P
19-JAN-2002: 2002US-0376522P
18-JAN-2002: 2002US-0376445P
17-JAN-2002: 2002US-0376368P
16-JAN-2002: 2002US-0376291P
15-JAN-2002: 2002US-0376214P
14-JAN-2002: 2002US-0376137P
13-JAN-2002: 200

PR 11-JUN-2002; 2002HS-0387634P

11-JUN-2002; 2002US-0387696P;
PR 11-JUN-2002; 2002US-0387696P;
ER 11-JUN-2002; 2002US-0387696P;
ER 11-JUN-2002; 2002US-0387696P;

11-JUN-2002; 2002US-0387836P

12-JUN-2002; 2002US-0387933P.

12-JUN-2002; 2002US-0387960P.

12-JUN-2002; 2002US-0388096P. PR

PR 14-JUN-2002; 2002US-0389118P.

PR 14-JUN-2002; 2002US-0389144P.

PR 17-JUN-2002; 2002US-0389729P.

PR 18-JUN-2002; 2002US-0389884P.

PR 19-JUN-2002; 2002US-0390209P.

PR 17-JUL-2002; 2002US-0396706P.

PR 09-AUG-2002; 2002US-0402156P.

PR 09-AUG-2002; 2002US-0402389P.

PR 12-AUG-2002; 2002US-0402816P.

PR 12-AUG-2002; 2002US-0402832P.

PR 13-AUG-2002; 2002US-0403459P.

PR 13-AUG-2002; 2002US-0403532P.

PR 13-AUG-2002; 2002US-0406317P.

```

PR 26-AUG-2002; 2002US-0406182P.
PR 26-AUG-2002; 2002US-0406355P.
PR 27-AUG-2002; 2002US-0406240P.
PR 12-SEP-2002; 2002US-0410084P.
PR 20-SEP-2002; 2002US-0412528P.
PR 23-SEP-2002; 2002US-0412731P.
PR 30-SEP-2002; 2002US-0418011P.
PR 30-SEP-2002; 2002US-041839P.
PR 30-SEP-2002; 2002US-041839P.
PR 30-SEP-2002; 2002US-041839P.
PR 09-OCT-2002; 2002US-041954P.
PR 09-OCT-2002; 2002US-041954P.
PR 09-OCT-2002; 2002US-041954P.
PR 23-OCT-2002; 2002US-0420639P.
PR 28-OCT-2002; 2002US-0421186P.
PR 31-OCT-2002; 2002US-0422690P.
PR 01-NOV-2002; 2002US-0423130P.
PR 05-NOV-2002; 2002US-0423798P.
PR 05-NOV-2002; 2002US-0423798P.
PR 12-NOV-2002; 2002US-0425453P.
XX
XX
XX (CURA-) CURAGEN CORP.
XX
XX
XX Alabrook JP, Alvarez E, Anderson DW, Boldog FL, Caeman SU;
XX Catterton E, Chapoval A, Crabtree-Bokor JR, Edinger SR, Ellerman K;
XX Ettenberg S, Gargoli EA, Gerlach VL, Gorman L, Gunther E, Guo X;
XX Gusev VY, Hermann JL, Ji W, Kekuda R, Li L, Liu X, MacDougall JR,
XX MacLachlan T, Malyancker UM, Mexick AJ, Miller I, Mishra VS;
XX Padigaru M, Patutajan M, Pena CE, Peyman JA, Raha D, Rastelli L;
XX Rieger DK, Rothenberg ME, Sciore P, Shenoy SG, Shmets RA;
XX Smithson G, Spytek KA, Stone DJ, Vernet CAM, Voss EZ, Zhong M;
XX Zhong H;
XX
XX WPI; 2004-081935/08.
XX P-PSDB; ADH72190.
XX
XX
XX New NOVX polypeptides and nucleic acid molecules useful for preventing or
XX treating NOVX-associated disorders, e.g. cancer, diabetes, infection or
XX obesity, and in chromosome mapping, tissue typing or pharmacogenomics.
XX
XX Example 51; SEQ ID NO 1085; 1880bp; English.
XX
XX
XX The invention relates to a novel isolated polypeptide (NOVX). A
XX polypeptide of the invention has cytosolic, immunomodulator,
XX neuroprotective, neurotropic, anorectic, antidiabetic, antimicrobial, and
XX antihypaemic activity, and may have a use in gene therapy, and as a
XX vaccine. The polypeptides are encoded by NOVX polynucleotides comprising
XX any of the 303 fully defined nucleotide sequences given in the
XX CC specification. The polypeptide is useful in the manufacture of a
XX CC medicament for treating a syndrome associated with a human disease. The
XX CC polypeptide, polynucleotide and antibody are useful in diagnosing,
XX CC treating or preventing NOVX-associated disorders, e.g. cancer, cachexia,
XX CC Alzheimer's disease, Parkinson's disease, obesity, diabetes, infectious
XX CC diseases, metabolic syndrome X or dyslipidaemias. The nucleic acids are
XX CC further used as hybridisation probes, in chromosome mapping, tissue
XX CC typing, preventive medicine, and pharmacogenomics. The present sequence
XX CC encodes a NOVX polypeptide of the invention.
XX
XX
XX Sequence 5101 BP; 1271 A; 1301 C; 1309 G; 1220 T; 0 U; 0 Other;
SQ
XX
XX Query Match 85.6%; Score 15.4; DB 12; Length 5101;
XX Best Local Similarity 94.1%; Pred. No. 1e+03;
XX Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 2 AGACACACAGATTGCG 18
XX |||||
DB 396 AGCACACACAGATTGCG 412
XX |||||

```

```

XX
XX 29-JAN-2004 (first entry)
XX
XX Human endometrial specific gene, SEQ ID NO 122.
DE
XX
XX cytosolic; vaccine; human; endometrial specific genes;
XX endometrial specific protein; endometrial cancer; ds.
XX
XX Homo sapiens.
OS
XX
XX MO2003060081-A2.
PN
XX
XX 24-JUL-2003.
PD
XX
XX 23-DEC-2002; 2002WO-US041612.
PF
XX
XX 21-DEC-2001; 2001US-0342756P.
PR
XX
XX (DIAD-) DIADEXUS INC.
PA
XX
XX Sun Y, Liu C;
PI
XX
XX WPI; 2003-577666/54.
XX
XX
XX Nucleic acid molecules and polypeptides useful for diagnosing and
XX treating endometrial cancer and non-cancerous disease states in
XX endometrial.
XX
XX Claim 1; SEQ ID NO 122; 824bp; English.
XX
XX
XX The invention comprises the amino acid and DNA sequences of human
XX endometrial specific genes and proteins. The DNA and protein sequences of
XX the invention are useful for diagnosing, imaging and treating a patient
XX with endometrial cancer. The present DNA sequence represents a human
XX endometrial specific gene of the invention.
XX
XX Sequence 5188 BP; 1196 A; 1390 C; 1471 G; 1131 T; 0 U; 0 Other;
SQ
XX
XX Query Match 85.6%; Score 15.4; DB 10; Length 5188;
XX Best Local Similarity 94.1%; Pred. No. 1e+03;
XX Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 2 AGACACACAGATTGCG 18
XX |||||
DB 877 AGCACACACAGATTGCG 893
XX |||||

```

```

RESULT 41
ADBE72688
ID ADE72688 standard; DNA; 5204 BP.
XX
XX ADE72688;
AC
XX
XX 29-JAN-2004 (first entry)
DT
XX
XX Human endometrial specific gene, SEQ ID NO 128.
DE
XX
XX cytosolic; vaccine; human; endometrial specific genes;
XX endometrial specific protein; endometrial cancer; ds.
KW
XX
XX Homo sapiens.
XX
XX MO2003060081-A2.
PN
XX
XX 24-JUL-2003.
PD
XX
XX 23-DEC-2002; 2002WO-US041612.
PF
XX
XX 21-DEC-2001; 2001US-0342756P.
PR
XX
XX (DIAD-) DIADEXUS INC.
PA
XX
XX Sun Y, Liu C;
PI

```

XX DBI; 2003-577666/54.
 XX Nucleic acid molecules and polypeptides useful for diagnosing and
 PT treating endometrial cancer and non-cancerous disease states in
 PT endometrial.
 XX
 PS Claim 1; SEQ ID NO 128; 824pp; English.
 CC The invention comprises the amino acid and DNA sequences of human
 CC endometrial specific genes and proteins. The DNA and protein sequences of
 CC the invention are useful for diagnosing, imaging and treating a patient
 CC with endometrial cancer. The present DNA sequence represents a human
 CC endometrial specific gene of the invention.
 XX
 SQ Sequence 5204 BP; 1171 A; 1436 C; 1476 G; 1120 T; 0 U; 1 Other;
 Query Match 85.6%; Score 15.4; DB 10; Length 5204;
 Best Local Similarity 94.1%; Pred. No. 1e+03;
 Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 2 AGAACAACAAGATTGCG 18
 DB 952 AGCAACAACAAGATTGCG 968
 RESULT 42
 ADE72669
 ID ADE72669 standard; DNA; 5263 BP.
 XX ADE72669;
 XX 29-JAN-2004 (first entry)
 XX Human endometrial specific gene, SEQ ID NO 109.
 DE cytostatic; vaccine; human; endometrial specific genes;
 KM endometrial specific protein; endometrial cancer; ds.
 XX Homo sapiens.
 OS
 XX WO2003060081-A2.
 XX 24-JUL-2003.
 XX 23-DEC-2002; 2002WO-US041612.
 XX 21-DEC-2001; 2001US-0342756P.
 XX (DIAD-) DIADEXUS INC.
 PA Sun Y, Liu C;
 PI WPI; 2003-577666/54.
 XX Nucleic acid molecules and polypeptides useful for diagnosing and
 PT treating endometrial cancer and non-cancerous disease states in
 PT endometrial.
 XX Claim 1; SEQ ID NO 109; 824pp; English.
 CC The invention comprises the amino acid and DNA sequences of human
 CC endometrial specific genes and proteins. The DNA and protein sequences of
 CC the invention are useful for diagnosing, imaging and treating a patient
 CC with endometrial cancer. The present DNA sequence represents a human
 CC endometrial specific gene of the invention.
 XX
 SQ Sequence 5263 BP; 1211 A; 1412 C; 1485 G; 1155 T; 0 U; 0 Other;
 Query Match 85.6%; Score 15.4; DB 10; Length 5263;
 Best Local Similarity 94.1%; Pred. No. 1e+03; 1; Indels 0; Gaps 0;
 Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 AGAACAACAAGATTGCG 18
 DB 952 AGCAACAACAAGATTGCG 968
 RESULT 43
 ADE72689
 ID ADE72689 standard; DNA; 5273 BP.
 XX ADE72689;
 XX 29-JAN-2004 (first entry)
 XX Human endometrial specific gene, SEQ ID NO 129.
 DE cytostatic; vaccine; human; endometrial specific genes;
 KM endometrial specific protein; endometrial cancer; ds.
 XX Homo sapiens.
 OS
 XX WO2003060081-A2.
 XX 24-JUL-2003.
 XX 23-DEC-2002; 2002WO-US041612.
 XX 21-DEC-2001; 2001US-0342756P.
 XX (DIAD-) DIADEXUS INC.
 PA Sun Y, Liu C;
 PI WPI; 2003-577666/54.
 XX Nucleic acid molecules and polypeptides useful for diagnosing and
 PT treating endometrial cancer and non-cancerous disease states in
 PT endometrial.
 XX Claim 1; SEQ ID NO 129; 824pp; English.
 PS The invention comprises the amino acid and DNA sequences of human
 CC endometrial specific genes and proteins. The DNA and protein sequences of
 CC the invention are useful for diagnosing, imaging and treating a patient
 CC with endometrial cancer. The present DNA sequence represents a human
 CC endometrial specific gene of the invention.
 XX
 SQ Sequence 5273 BP; 1216 A; 1419 C; 1482 G; 1156 T; 0 U; 0 Other;
 Query Match 85.6%; Score 15.4; DB 10; Length 5273;
 Best Local Similarity 94.1%; Pred. No. 1e+03; 1; Indels 0; Gaps 0;
 Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 2 AGAACAACAAGATTGCG 18
 DB 952 AGCAACAACAAGATTGCG 968
 RESULT 44
 ADN05928
 ID ADN05928 standard; cDNA; 5287 BP.
 XX ADN05928;
 XX 01-JUL-2004 (first entry)
 XX Antiperoicatic cDNA sequence #1198.
 DE ds; gene; antiperoicatic; gene therapy; psoriasis; diagnosis.
 XX Homo sapiens.
 OS
 XX WO2004028479-A2.

PD 08-APR-2004.
XX
XX 25-SEP-2003; 2003WO-US030907.
XX
XX 25-SEP-2002; 2002US-0414006P.
XX
XX (GETH) GENENTECH INC.
XX
XX Bodary S, Clark H, Jackman J, Schenfeld J, Williams PM, Wood WT,
PI Wu TD;
XX WPI; 2004-305105/28.
XX P-PSDB; ADN05929.
XX
XX New PRO nucleic acid or polypeptide, useful for preparing a
PT pharmaceutical composition for diagnosing or treating psoriasis in a
PT mammal.
XX
XX Claim 1; SEQ ID NO 2323; 3069pp; English.
XX
XX The invention relates to novel polynucleotide and polypeptides for
CC treating psoriasis or a sequence having at least 80% identity to the
CC above sequences. The nucleic acid is useful for preparing a composition
CC for diagnosing or treating psoriasis in a mammal. This sequence
CC corresponds to one of the polynucleotides of the invention.
XX
SQ Sequence 5287 BP; 1245 A; 1411 C; 1469 G; 1162 T; 0 U; 0 Other;
Query Match 85.6%; Score 15.4; DB 12; Length 5287;
Best Local Similarity 94.1%; Pred. No. 1e+03; 1; Indels 0; Gaps 0;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 2 AGAACACACAGATTGCG 18
DB 887 AGCACACACAGATTGCG 903
RESULT 45
ADE72676
ID ADE72676 standard; DNA; 5332 BP.
XX
XX ADE72676;
XX
XX 29-JAN-2004 (first entry)
XX
XX Human endometrial specific gene, SEQ ID NO 116.
DE
XX cytosatic; vaccine; human; endometrial specific genes;
KW endometrial specific protein; endometrial cancer; ds.
XX
XX Homo sapiens.
OS
XX
XX WO2003060081-A2.
XX
XX 24-JUL-2003.
XX
XX 23-DEC-2002; 2002WO-US041612.
XX
XX 21-DEC-2001; 2001US-0342756P.
XX
XX (DIAD-) DIADEXUS INC.
XX
XX Sun Y, Liu C;
XX WPI; 2003-577666/54.
XX
XX Nucleic acid molecules and polypeptides useful for diagnosing and
PT treating endometrial cancer and non-cancerous disease states in
PT endometrial.
XX
XX Claim 1; SEQ ID NO 116; 824pp; English.
XX
XX The invention comprises the amino acid and DNA sequences of human

CC endometrial specific genes and proteins. The DNA and protein sequences of
CC the invention are useful for diagnosing, imaging and treating a patient
CC with endometrial cancer. The present DNA sequence represents a human
CC endometrial specific gene of the invention.
XX
SQ Sequence 5332 BP; 1232 A; 1434 C; 1499 G; 1167 T; 0 U; 0 Other;
Query Match 85.6%; Score 15.4; DB 10; Length 5332;
Best Local Similarity 94.1%; Pred. No. 1e+03; 1; Indels 0; Gaps 0;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 2 AGAACACACAGATTGCG 18
DB 952 AGCACACACAGATTGCG 968
RESULT 46
ADE72666
ID ADE72666 standard; DNA; 5337 BP.
XX
XX ADE72666;
XX
XX 29-JAN-2004 (first entry)
XX
XX Human endometrial specific gene, SEQ ID NO 106.
DE
XX cytosatic; vaccine; human; endometrial specific genes;
KW endometrial specific protein; endometrial cancer; ds.
XX
XX Homo sapiens.
OS
XX
XX WO2003060081-A2.
XX
XX 24-JUL-2003.
XX
XX 23-DEC-2002; 2002WO-US041612.
XX
XX 21-DEC-2001; 2001US-0342756P.
XX
XX (DIAD-) DIADEXUS INC.
XX
XX Sun Y, Liu C;
XX WPI; 2003-577666/54.
XX
XX Nucleic acid molecules and polypeptides useful for diagnosing and
PT treating endometrial cancer and non-cancerous disease states in
PT endometrial.
XX
XX Claim 1; SEQ ID NO 106; 824pp; English.
XX
XX The invention comprises the amino acid and DNA sequences of human
CC endometrial specific genes and proteins. The DNA and protein sequences of
CC the invention are useful for diagnosing, imaging and treating a patient
CC with endometrial cancer. The present DNA sequence represents a human
CC endometrial specific gene of the invention.
XX
SQ Sequence 5337 BP; 1243 A; 1432 C; 1495 G; 1167 T; 0 U; 0 Other;
Query Match 85.6%; Score 15.4; DB 10; Length 5337;
Best Local Similarity 94.1%; Pred. No. 1e+03; 1; Indels 0; Gaps 0;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 2 AGAACACACAGATTGCG 18
DB 948 AGCACACACAGATTGCG 964
RESULT 47
ADE72667
ID ADE72667 standard; DNA; 5353 BP.
XX
XX ADE72667;
AC

```
XX 23-JAN-2004 (first entry)
DT Human endometrial specific gene, SEQ ID NO 107.
XX
XX cytosstatic; vaccine; human; endometrial specific genes;
KM endometrial specific protein; endometrial cancer; ds.
XX
XX Homo sapiens.
OS
XX WO2003060081-A2.
PN
XX 24-JUL-2003.
PD
XX 23-DEC-2002; 2002WO-US041612.
PF
XX 21-DEC-2001; 2001US-0342756P.
PR
XX (DIAD-) DIADEXUS INC.
PA
XX Sun Y, Liu C;
PI WPI; 2003-577666/54.
XX
XX Nucleic acid molecules and polypeptides useful for diagnosing and
PT treating endometrial cancer and non-cancerous disease states in
PT endometrial.
XX
XX Claim 1; SEQ ID NO 107; 824pp; English.
XX
XX The invention comprises the amino acid and DNA sequences of human
CC endometrial specific genes and proteins. The DNA and protein sequences of
CC the invention are useful for diagnosing, imaging and treating a patient
CC with endometrial cancer. The present DNA sequence represents a human
CC endometrial specific gene of the invention.
XX
XX Sequence 5353 BP; 1245 A; 1434 C; 1499 G; 1175 T; 0 U; 0 Other;
SQ
XX
XX Query Match 85.6%; Score 15.4; DB 10; Length 5353;
XX Best Local Similarity 94.1%; Pred. No. 1e+03;
XX Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
OY 2 AGAACAACAGATTGCG 18
DB 952 AGCACAACAGATTGCG 968
XX
XX RESULT 48
XX ADE72681
XX ID ADE72681 standard; DNA; 5639 BP.
XX
XX ADE72681;
XX
XX 29-JAN-2004 (first entry)
DT
XX Human endometrial specific gene, SEQ ID NO 121.
DE
XX cytosstatic; vaccine; human; endometrial specific genes;
XX endometrial specific protein; endometrial cancer; ds.
XX
XX Homo sapiens.
OS
XX WO2003060081-A2.
PN
XX 24-JUL-2003.
PD
XX 23-DEC-2002; 2002WO-US041612.
PF
XX 21-DEC-2001; 2001US-0342756P.
PR
XX (DIAD-) DIADEXUS INC.
PA
XX Sun Y, Liu C;
PI
```

```
XX WPI; 2003-577666/54.
DR
XX Nucleic acid molecules and polypeptides useful for diagnosing and
PT treating endometrial cancer and non-cancerous disease states in
PT endometrial.
XX
XX Claim 1; SEQ ID NO 121; 824pp; English.
XX
XX The invention comprises the amino acid and DNA sequences of human
CC endometrial specific genes and proteins. The DNA and protein sequences of
CC the invention are useful for diagnosing, imaging and treating a patient
CC with endometrial cancer. The present DNA sequence represents a human
CC endometrial specific gene of the invention.
XX
XX Sequence 5639 BP; 1273 A; 1495 C; 1638 G; 1233 T; 0 U; 0 Other;
SQ
XX
XX Query Match 85.6%; Score 15.4; DB 10; Length 5639;
XX Best Local Similarity 94.1%; Pred. No. 1.1e+03;
XX Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
OY 2 AGAACAACAGATTGCG 18
DB 1259 AGCACAACAGATTGCG 1275
XX
XX RESULT 49
XX ADE72685
XX ID ADE72685 standard; DNA; 5684 BP.
XX
XX ADE72685;
XX
XX 29-JAN-2004 (first entry)
DT
XX Human endometrial specific gene, SEQ ID NO 125.
DE
XX cytosstatic; vaccine; human; endometrial specific genes;
XX endometrial specific protein; endometrial cancer; ds.
XX
XX Homo sapiens.
OS
XX WO2003060081-A2.
PN
XX 24-JUL-2003.
PD
XX 23-DEC-2002; 2002WO-US041612.
PF
XX 21-DEC-2001; 2001US-0342756P.
PR
XX (DIAD-) DIADEXUS INC.
PA
XX Sun Y, Liu C;
PI WPI; 2003-577666/54.
XX
XX Nucleic acid molecules and polypeptides useful for diagnosing and
PT treating endometrial cancer and non-cancerous disease states in
PT endometrial.
XX
XX Claim 1; SEQ ID NO 125; 824pp; English.
XX
XX The invention comprises the amino acid and DNA sequences of human
CC endometrial specific genes and proteins. The DNA and protein sequences of
CC the invention are useful for diagnosing, imaging and treating a patient
CC with endometrial cancer. The present DNA sequence represents a human
CC endometrial specific gene of the invention.
XX
XX Sequence 5684 BP; 1399 A; 1497 C; 1594 G; 1194 T; 0 U; 0 Other;
SQ
XX
XX Query Match 85.6%; Score 15.4; DB 10; Length 5684;
XX Best Local Similarity 94.1%; Pred. No. 1.1e+03;
XX Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

OY 2 AGACACACAGATTGGC 18
 |||||
 DB 952 AGACACACAGATTGGC 968

RESULT 50

ADE72670
 ID ADE72670 standard; DNA; 5857 BP.

AC ADE72670;

DT 29-JAN-2004 (first entry)

DE Human endometrial specific gene, SEQ ID NO 110.

KW cytotatic; vaccine; human; endometrial specific genes;
 endometrial specific protein; endometrial cancer; ds.

OS Homo sapiens.

PN WO2003060081-A2.

PD 24-JUL-2003.

PF 23-DEC-2002; 2002WO-US041612.

PR 21-DEC-2001; 2001US-0342756P.

PA (DIAD-) DIADEXUS INC.

PI Sun Y, Liu C;

DR WPI; 2003-577666/54.

PT Nucleic acid molecules and polypeptides useful for diagnosing and
 treating endometrial cancer and non-cancerous disease states in
 endometrial.

PS Claim 1; SEQ ID NO 110; 824bp; English.

CC The invention comprises the amino acid and DNA sequences of human
 endometrial specific genes and proteins. The DNA and protein sequences of
 the invention are useful for diagnosing, imaging and treating a patient
 with endometrial cancer. The present DNA sequence represents a human
 endometrial specific gene of the invention.

CC Sequence 5857 BP; 1368 A; 1567 C; 1600 G; 1322 T; 0 U; 0 Other;

Query Match 85.6%; Score 15.4; DB 10; Length 5857;

Best Local Similarity 94.1%; Pred. No. 1.1e+03;

Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 AGACACACAGATTGGC 18
 |||||
 DB 952 AGACACACAGATTGGC 968

Search completed: December 3, 2004, 02:24:15
 Job time : 291.211 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Comphen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 2, 2004, 23:31:20 ; Search time 563.321 Seconds
(without alignments)
1509.457 Million cell updates/sec

Title: US-10-050-189a-6
Perfect score: 18
Sequence: 1 gagaaacacaaagatcgc 18

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4526729 seqs, 23644849745 residues
Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 1000 summaries

Database :
1: GenEmbl:*
2: gb_bac:*
3: gb_mtg:*
4: gb_in:*
5: gb_cm:*
6: gb_cv:*
7: gb_pat:*
8: gb_ph:*
9: gb_pl:*
10: gb_pt:*
11: gb_ro:*
12: gb_sts:*
13: gb_sy:*
14: gb_un:*
15: gb_vi:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	17	94.4	110000	8	CR382126_06
2	16.4	91.1	10610	1	AE006452
3	16.4	91.1	39228	1	MLCB1788
4	16.4	91.1	100867	8	AC008047
5	16.4	91.1	110000	2	LMFLCHR36_27
6	16.4	91.1	124215	2	AC143118
7	16.4	91.1	142018	8	AY146587
8	16.4	91.1	169261	2	CR352284
9	16.4	91.1	214536	2	CR382377
10	16.4	91.1	239570	2	AC111804
11	16.4	91.1	256123	2	AC111825
12	16.4	91.1	342300	1	MLCPRN8
13	16	88.9	19	6	AX481359
14	16	88.9	64167	2	AC009112
15	16	88.9	66479	6	AX676048
16	16	88.9	78376	9	AL359692
17	16	88.9	88701	8	AC135161
18	16	86.9	110000	8	CR382130_35
19	16	88.9	110000	8	CR382133_06

20	16	88.9	154732	9	AC011700	AC011700 Homo sapi
21	16	88.9	157381	2	AC129818	AC129818 Rattus no
22	16	88.9	168612	2	AC007341	AC007341 Homo sapi
23	16	88.9	180837	9	AC007333	AC007333 Homo sapi
24	16	88.9	214026	2	AC111389	AC111389 Rattus no
25	16	88.9	236172	2	AC094494	AC094494 Rattus no
26	16	88.9	251032	2	AC126106	AC126106 Rattus no
27	16	88.9	292516	2	AC094425	AC094425 Rattus no
28	15.4	85.6	485	6	AX340908	AX340908 Sequence
29	15.4	85.6	534	6	AX345508	AX345508 Sequence
30	15.4	85.6	886	5	AX344560	AX344560 Sinogast
31	15.4	85.6	887	5	AX344557	AX344557 Sinogast
32	15.4	85.6	887	5	AX344559	AX344559 Sinogast
33	15.4	85.6	887	5	AX344561	AX344561 Sinogast
34	15.4	85.6	887	5	AX344562	AX344562 Sinogast
35	15.4	85.6	888	5	AX344558	AX344558 Sinogast
36	15.4	85.6	1524	14	JSR1TMPB	JSR1TMPB
37	15.4	85.6	1664	8	AK109670	AK109670 Oryza sat
38	15.4	85.6	1950	8	AY692867	AY692867 Saccharom
39	15.4	85.6	1997	8	AB091267	AB091267 Marchanti
40	15.4	85.6	2186	6	BD095617	BD095617 Process f
41	15.4	85.6	2945	6	CO716675	CO716675 Sequence
42	15.4	85.6	3178	6	AR336033	AR336033 Sequence
43	15.4	85.6	3178	6	AR201081	AR201081 Sequence
44	15.4	85.6	3178	6	BD007547	BD007547 Process o
45	15.4	85.6	3178	6	BD009176	BD009176 Method fo
46	15.4	85.6	3178	6	SCGT2A	X71660 S.cerevisia
47	15.4	85.6	3655	1	ECOLYSP	M89774 Escherichia
48	15.4	85.6	3831	2	AB091102	AB091102 Marchanti
49	15.4	85.6	3831	2	AB091102	AB091102 Marchanti
50	15.4	85.6	3831	2	AB091102	AB091102 Marchanti
51	15.4	85.6	4717	6	AX463544	AX463544 Sequence
52	15.4	85.6	4759	6	AX463530	AX463530 Sequence
53	15.4	85.6	4762	9	AF381545	AF381545 Homo sapi
54	15.4	85.6	5044	3	AF293351	AF293351 Anopheles
55	15.4	85.6	5287	9	AB050468	AB050468 Homo sapi
56	15.4	85.6	5303	9	BC071561	BC071561 Homo sapi
57	15.4	85.6	6696	4	AF136225	AF136225 Ovis arie
58	15.4	85.6	7043	6	BD193694	BD193694 Enterococ
59	15.4	85.6	7455	14	AF105220	AF105220 Ovine pul
60	15.4	85.6	7462	14	A27950	A27950 JSRV genom
61	15.4	85.6	7940	4	JSRCG	M80216 Jaagsiekte
62	15.4	85.6	8924	1	AF153615	AF153615 Ovis arie
63	15.4	85.6	8924	1	AB014911	AB014911 Streptococ
64	15.4	85.6	10182	1	AB012849	AB012849 Chlorella
65	15.4	85.6	10472	1	AB015240	AB015240 Shigella
66	15.4	85.6	10474	1	AE005448	AE005448 Escherich
67	15.4	85.6	11791	14	AF357971	AF357971 Ovine pul
68	15.4	85.6	12905	8	AF353329	AF353329 Kallisto
69	15.4	85.6	23223	2	AC005975	AC005975 Drosophila
70	15.4	85.6	28586	3	CEC14465	Z81470 Caenorhabdi
71	15.4	85.6	30501	3	CEC10E12	Z29561 Caenorhabdi
72	15.4	85.6	35412	8	SPC188	AL049662 S.pombe c
73	15.4	85.6	37170	3	CEFC30A10	Z81072 Caenorhabdi
74	15.4	85.6	41803	8	SPC584	AL132824 S.pombe c
75	15.4	85.6	57745	3	CEY6782A	AL132951 Caenorhab
76	15.4	85.6	62355	2	AC084472	AC084472 Caenorhab
77	15.4	85.6	62588	2	AC101236	AC101236 Mus muscu
78	15.4	85.6	75317	8	SCA357	AC021389 Homo sapi
79	15.4	85.6	75888	1	ECOHU47	Z80859 S.cerevisia
80	15.4	85.6	75898	14	AY190604	U00007 47 to 48 ce
81	15.4	85.6	77670	7	AX050604	AY190604 Halovirus
82	15.4	85.6	77670	7	AX050604	AY190604 Halovirus
83	15.4	85.6	79459	9	AB022207	AF222060 Halovirus
84	15.4	85.6	92602	9	AC010550	AB022217 Arabidops
85	15.4	85.6	99661	2	AC137530	AC010550 Homo sapi
86	15.4	85.6	110000	1	BCX908798_22	AC137530 Takifugu
87	15.4	85.6	110000	2	AC129474_1	Continuation (23 o
88	15.4	85.6	110000	2	AC129474_1	Continuation (23 o
89	15.4	85.6	110000	2	AC138525_0	Continuation (12 o
90	15.4	85.6	110000	2	AC138525_0	Continuation (12 o
91	15.4	85.6	110000	2	CR382128_27	Continuation (12 o
92	15.4	85.6	113566	8	AC002330	Continuation (12 o

93	15.4	85.6	116265	5	AC140941	AC140941 Gallus ga	c 166	15	83.3	109431	9	AC016696	AC016696 Homo sapi	
94	15.4	85.6	123847	8	AC138002	AC138002 Oryza sat	c 167	15	83.3	110000	8	CR382125	Continuance (9 of	
95	15.4	85.6	127299	8	AC147434	AC147434 Medicago	c 168	15	83.3	112328	8	AL390997	Human DNA	
96	15.4	85.6	127300	8	AC147429	AC147429 Medicago	c 169	15	83.3	147706	8	AC083945	AC083945 Oryza sat	
97	15.4	85.6	133116	2	AC031985	AC031985 Homo sapi	c 170	15	83.3	170051	8	AP005772	AP005772 Oryza sat	
98	15.4	85.6	139215	8	AC126780	AC126780 Medicago	c 171	15	83.3	171617	2	AC013616	AC013616 Homo sapi	
99	15.4	85.6	146632	2	AC115935	AC115935 Mus muscu	c 172	15	83.3	173223	2	AP003377	AP003377 Homo sapi	
100	15.4	85.6	146775	2	AC013153	AC013153 Drosophil	c 173	15	83.3	173759	2	CR450686	CR450686 Dantio rer	
101	15.4	85.6	150339	2	AC141750	AC141750 Apis mell	c 174	15	83.3	181460	10	AL672181	AL672181 Mouse DNA	
102	15.4	85.6	152836	2	CR450718	CR450718 Dantio rer	c 175	15	83.3	181952	2	AC107053	AC107053 Homo sapi	
103	15.4	85.6	164176	2	AC116409	AC116409 Mus muscu	c 176	15	83.3	183342	9	AL450345	AL450345 Human DNA	
104	15.4	85.6	167627	2	AC079382	AC079382 Homo sapi	c 177	15	83.3	187265	8	AP004836	AP004836 Oryza sat	
105	15.4	85.6	174151	2	AC011182	AC011182 Homo sapi	c 178	15	83.3	187688	9	AP003366	AP003366 Homo sapi	
106	15.4	85.6	174691	9	AC018563	AC018563 Homo sapi	c 179	15	83.3	192128	2	AC013382	AC013382 Homo sapi	
107	15.4	85.6	175448	3	AC099035	AC099035 Drosophil	c 180	15	83.3	195835	2	AC107840	AC107840 Mus muscu	
108	15.4	85.6	181463	3	AC008352	AC008352 Drosophil	c 181	15	83.3	201210	2	AC024932	AC024932 Homo sapi	
109	15.4	85.6	186415	9	AC008675	AC008675 Homo sapi	c 182	15	83.3	203741	2	AC128206	AC128206 Rattus no	
110	15.4	85.6	186418	9	AC018757	AC018757 Homo sapi	c 183	15	83.3	211855	2	AC106459	AC106459 Rattus no	
111	15.4	85.6	189037	2	AC150026	AC150026 Papio anu	c 184	15	83.3	216402	2	AC099109	AC099109 Rattus no	
112	15.4	85.6	189724	2	CR383675	CR383675 Dantio rer	c 185	15	83.3	216680	2	AC021211	AC021211 Homo sapi	
113	15.4	85.6	189827	5	BR890610	BR890610 Zebrafish	c 186	15	83.3	222361	2	AC110091	AC110091 Mus muscu	
114	15.4	85.6	192063	9	AC099665	AC099665 Homo sapi	c 187	15	83.3	229315	2	AC105651	AC105651 Rattus no	
115	15.4	85.6	196491	9	AC092832	AC092832 Homo sapi	c 188	15	83.3	233157	2	AC105468	AC105468 Rattus no	
116	15.4	85.6	199615	8	ATCNR1V7	AL161495 Arabidops	c 189	15	83.3	235045	2	AC119588	AC119588 Rattus no	
117	15.4	85.6	203107	2	AC139255	AC139255 Homo sapi	c 190	15	83.3	237080	2	AC106251	AC106251 Rattus no	
118	15.4	85.6	205764	2	AC079478	AC079478 Mus muscu	c 191	15	83.3	238597	2	AC131857	AC131857 Rattus no	
119	15.4	85.6	213514	2	AC079500	AC079500 Mus muscu	c 192	15	83.3	238722	2	AC098108	AC098108 Rattus no	
120	15.4	85.6	213514	2	AC079500	AC079500 Mus muscu	c 193	15	83.3	239760	2	AC126206	AC126206 Rattus no	
121	15.4	85.6	219429	10	AL671925	AL671925 Mouse DNA	c 194	15	83.3	240479	2	AC096377	AC096377 Rattus no	
122	15.4	85.6	222629	2	AC125802	AC125802 Rattus no	c 195	15	83.3	274786	2	AC109661	AC109661 Rattus no	
123	15.4	85.6	225158	5	AL953901	AL953901 Zebrafish	c 196	15	83.3	300029	8	AE017083	AE017083 Oryza sat	
124	15.4	85.6	225654	5	AL953842	AL953842 Zebrafish	c 197	15	83.3	348402	2	AC120619	AC120619 Rattus no	
125	15.4	85.6	230188	2	AC091604	AC091604 Mus muscu	c 198	14.8	82.2	294	11	G71609	G71609 A62987734FW	
126	15.4	85.6	236612	2	AC111126	AC111126 Mus muscu	c 199	14.8	82.2	304	11	G71923	G71923 A6496534FW	
127	15.4	85.6	245071	2	AC134078	AC134078 Rattus no	c 200	14.8	82.2	320	11	G70842	G70842 A62987734FB	
128	15.4	85.6	245558	2	AC079514	AC079514 Mus muscu	c 201	14.8	82.2	400	11	BV143261	BV143261 Bos tauru	
129	15.4	85.6	249381	2	AC009937	AC009937 Homo sapi	c 202	14.8	82.2	467	4	AY138588	AY138588 Bos tauru	
130	15.4	85.6	254250	1	AP005350	AP005350 Vibrio vu	c 203	14.8	82.2	467	6	AX658805	AX658805 Sequence	
131	15.4	85.6	265053	1	AC098057	AC098057 Rattus no	c 204	14.8	82.2	468	11	BV143259	BV143259 P203027 Z	
132	15.4	85.6	268857	1	AP002560	AP002560 Escherich	c 205	14.8	82.2	471	11	BV143262	BV143262 P203027 Z	
133	15.4	85.6	279691	2	AC129258	AC129258 Rattus no	c 206	14.8	82.2	481	11	BV143265	BV143265 P203027 Z	
134	15.4	85.6	283700	3	AE003459	AE003459 Drosophill	c 207	14.8	82.2	483	11	BV143267	BV143267 P203027 Z	
135	15.4	85.6	295689	1	AE016985	AE016985 Shigella	c 208	14.8	82.2	488	11	BV143268	BV143268 P203027 Z	
136	15.4	85.6	300029	1	AE016952	AE016952 Enterococ	c 209	14.8	82.2	488	11	BV143260	BV143260 P203027 Z	
137	15.4	85.6	300150	1	AP004598	AP004598 Oceanobac	c 210	14.8	82.2	489	11	BV143256	BV143256 P203027 Z	
138	15.4	85.6	300964	8	AE016890	AE016890 Eremothec	c 211	14.8	82.2	491	6	BD178253	BD178253 Mold poly	
139	15.4	85.6	301130	1	AE016763	AE016763 Escherich	c 212	14.8	82.2	491	6	BD179332	BD179332 Polynucleo	
140	15.4	85.6	301332	1	AE017012	AE017012 Bacillus	c 213	14.8	82.2	491	11	BV143270	BV143270 P203027 Z	
141	15.4	85.6	306803	2	AE017161	AE017161 Prochloro	c 214	14.8	82.2	493	11	BV143264	BV143264 P203027 Z	
142	15.4	85.6	312163	2	AC126132	AC126132 Rattus no	c 215	14.8	82.2	493	11	BV143265	BV143265 P203027 Z	
143	15.4	85.6	313573	3	CEY57611C	CEY57611C Caenorhabd	c 216	14.8	82.2	493	11	BV143267	BV143267 P203027 Z	
144	15.4	85.6	319485	2	AC078979	AC078979 Homo sapi	c 217	14.8	82.2	495	11	BV143268	BV143268 P203027 Z	
145	15.4	85.6	330724	2	CEY67H2	AL022475 Caenorhab	c 218	14.8	82.2	495	11	BV143271	BV143271 P203027 Z	
146	15.4	85.6	330724	2	CEY67H2	BD252112 Compositi	c 219	14.8	82.2	496	11	BV143269	BV143269 P203027 Z	
147	15.4	85.6	333	485	6	BD252112	BR360897 Sequence	c 220	14.8	82.2	569	11	BV000188	BV000188 Homo sapi
148	15.4	85.6	333	737	8	AR360897	BT000168 Arabidops	c 221	14.8	82.2	578	8	AY299269	AY299269 Arabidops
149	15.4	85.6	333	777	8	BT000168	BT000168 Arabidops	c 222	14.8	82.2	581	3	DMVATP14	DMVATP14 D.melanog
150	15.4	85.6	333	744	11	PM2G9G	AY163638 Arabidops	c 223	14.8	82.2	585	11	G99063	G99063 S208P6545RE
151	15.4	85.6	333	889	8	AY163638	AY087026 Arabidops	c 224	14.8	82.2	602	6	CO587817	CO587817 Sequence
152	15.4	85.6	333	891	8	AY087026	AX488833 Sequence	c 225	14.8	82.2	634	6	AX721295	AX721295 Sequence
153	15.4	85.6	333	1455	6	AX488833	Y09921 P. anomala U	c 226	14.8	82.2	700	14	HAUD2898	HAUD2898 Human adeno
154	15.4	85.6	333	1689	9	PAURA3	Z70288 Scenedesmus	c 227	14.8	82.2	722	8	AY299268	AY299268 Arabidops
155	15.4	85.6	333	36798	9	HSE78G1	AF204057 Scenedes	c 228	14.8	82.2	759	11	BV036917	BV036917 Arabidops
156	15.4	85.6	333	42919	8	AF204057	AB016876 Arabidops	c 229	14.8	82.2	813	11	BV062894	BV062894 Arabidops
157	15.4	85.6	333	44499	8	AB016876	AC090633 Homo sapi	c 230	14.8	82.2	840	8	BT012634	BT012634 Arabidops
158	15.4	85.6	333	66240	2	AC090633	AC090633 Arabidops	c 231	14.8	82.2	952	8	ABI26868	ABI26868 Aspergill
159	15.4	85.6	333	81609	8	AC027035	AC027035 Arabidops	c 232	14.8	82.2	993	8	GN132218	GN132218 Gnetum gn
160	15.4	85.6	333	84162	6	T22J18	AX652128 Sequence	c 233	14.8	82.2	1143	10	AF155474	AF155474 Sigmondon
161	15.4	85.6	333	95769	6	AX652128	AC013440 Genomic s	c 234	14.8	82.2	1143	10	AF296188	AF296188 Sigmondon
162	15.4	85.6	333	95769	6	AC013440	AC051647 Arabidops	c 235	14.8	82.2	1143	10	AF425215	AF425215 Sigmondon
163	15.4	85.6	333	97242	8	AC051630	AC016447 Arabidops	c 236	14.8	82.2	1143	10	AF425216	AF425216 Sigmondon
164	15.4	85.6	333	107816	8	AC016447	AC096717 Homo sapi	c 237	14.8	82.2	1143	10	AF425217	AF425217 Sigmondon
165	15.4	85.6	333	109027	9	AC096717		c 238	14.8	82.2	1202	8	AY485263	AY485263 Zea mays

239	14.8	82.2	1253	8	AY452809	AY452809 Fusarium	C 312	14.8	82.2	17716	3	U20862	U20862 Caenorhabdi
C 240	14.8	82.2	1292	8	BT236854	BT236854 Bos tauru	313	14.8	82.2	20704	2	LMFCHR5_7	Continuation (8 of
241	14.8	82.2	1336	8	AF212615	AF212615 Fusarium	314	14.8	82.2	20722	1	AE008869	AE008869 Salmonell
242	14.8	82.2	1336	8	AF212616	AF212616 Fusarium	C 315	14.8	82.2	23424	2	AC017219	AC017219 Drosophi
243	14.8	82.2	1336	8	AF212617	AF212617 Fusarium	C 316	14.8	82.2	24186	3	CER05H5	Z48795 Caenorhabdi
244	14.8	82.2	1336	8	AF212618	AF212618 Fusarium	C 317	14.8	82.2	24186	3	CER05H5	Z81542 Caenorhabdi
C 245	14.8	82.2	1359	6	AX078926	AX078926 Sequence	318	14.8	82.2	29216	2	AC136556	AC136556 Rattus no
C 246	14.8	82.2	1377	4	AF036953	AF036953 Felis cat	C 319	14.8	82.2	30352	3	U28412	U28412 Caenorhabdi
C 247	14.8	82.2	1480	8	AY084937	AY084937 Arabidops	320	14.8	82.2	35708	6	AX084516	AX084516 Sequence
C 248	14.8	82.2	1517	8	AK109902	AK109902 Oryza sat	321	14.8	82.2	35724	6	AX084519	AX084519 Sequence
C 249	14.8	82.2	1552	8	ATU82203	U82203 Arabidops	322	14.8	82.2	36114	6	AY183926	AY183926 Drosophi
C 250	14.8	82.2	1564	8	AY096554	AY096554 Arabidops	C 323	14.8	82.2	39165	3	AT183926	AT183926 Caenorhabdi
251	14.8	82.2	1692	8	AY224467	AY224467 Oryza sat	C 324	14.8	82.2	39235	3	AF016427	AF016427 Caenorhab
252	14.8	82.2	1821	4	BOVSL37B	L11063 Bos tauru	325	14.8	82.2	39694	2	AC144439	AC144439 Streptomy
C 253	14.8	82.2	1830	4	BOVSL37B	BOVSL37B Homo sapi	326	14.8	82.2	41097	1	AF016585	AF016585 Homo sapi
254	14.8	82.2	1863	8	AY065415	AY065415 Arabidops	C 327	14.8	82.2	41396	3	HSU73633	HSU73633 Caenorhabdi
C 255	14.8	82.2	2000	6	AX655526	AX655526 Sequence	C 328	14.8	82.2	42845	9	HSU73633	HSU73633 Caenorhabdi
C 256	14.8	82.2	2139	3	AY058544	AY058544 Drosophi	C 329	14.8	82.2	43721	3	U88315	U88315 Caenorhabdi
C 257	14.8	82.2	2157	6	CQ588615	CQ588615 Sequence	330	14.8	82.2	44380	1	AE001137	AE001137 Borrelia
C 258	14.8	82.2	2401	8	AK065899	AK065899 Oryza sat	331	14.8	82.2	46264	2	AC137490	AC137490 Homo sapi
C 259	14.8	82.2	2529	6	CQ804028	CQ804028 Sequence	C 332	14.8	82.2	47000	8	AP005896	AP005896 Oryza sat
C 260	14.8	82.2	2560	6	AK111850	AK111850 Oryza sat	C 333	14.8	82.2	47452	3	CER045N09	CER045N09 Caenorhab
261	14.8	82.2	2602	6	CQ587816	CQ587816 Sequence	334	14.8	82.2	48505	2	AC137489	AC137489 Homo sapi
C 262	14.8	82.2	2645	14	AY144570	AY144570 Rice blac	C 335	14.8	82.2	50641	8	AB008270	AB008270 Homo sapi
C 263	14.8	82.2	2645	14	RBL409148	AY144570 Rice blac	C 336	14.8	82.2	53797	9	AP001415	AP001415 Homo sapi
C 264	14.8	82.2	2727	6	AR320492	AR320492 Sequence	C 337	14.8	82.2	53905	9	AL359292	AL359292 Human DNA
C 265	14.8	82.2	2821	8	AF233592	AF233592 Arabidops	C 338	14.8	82.2	53981	9	CR388122	CR388122 Zebrafish
C 266	14.8	82.2	2859	8	AK099981	AK099981 Oryza sat	339	14.8	82.2	60729	2	AC100232	AC100232 Mus muscu
C 267	14.8	82.2	3030	6	AC014405	AC014405 Drosophi	C 340	14.8	82.2	66241	2	AC019757	AC019757 Drosophi
268	14.8	82.2	3081	6	AK777288	AK777288 Sequence	341	14.8	82.2	72828	2	AC026849	AC026849 Homo sapi
269	14.8	82.2	3098	6	CQ589908	CQ589908 Sequence	C 342	14.8	82.2	73617	3	AC018226	AC018226 Drosophi
270	14.8	82.2	3141	3	DMDWNT3MR	AX47736 D. melanog	C 343	14.8	82.2	74901	3	AC005474	AC005474 Homo sapi
271	14.8	82.2	3169	3	AF188994	AF188994 Arabidops	344	14.8	82.2	77939	8	AB005246	AB005246 Human DNA
272	14.8	82.2	3181	3	AF289998	AF289998 Drosophi	345	14.8	82.2	79303	9	AL592302	AL592302 Homo sapi
273	14.8	82.2	3213	3	AY122223	AY122223 Drosophi	C 346	14.8	82.2	79516	9	AC004834	AC004834 Lotus cor
274	14.8	82.2	3231	10	AK120280	AK120280 Oryza sat	C 347	14.8	82.2	80088	8	ATT31P16	ATT31P16 Homo sapi
C 275	14.8	82.2	3231	10	MMU487561	AX487521 Mus muscu	C 348	14.8	82.2	82389	8	AP004909	AP004909 Homo sapi
276	14.8	82.2	3346	1	MTDNAHEVA	AX487521 Mus muscu	C 349	14.8	82.2	82406	9	AC113430	AC113430 Homo sapi
277	14.8	82.2	3346	10	MMU487960	AX487521 Mus muscu	C 350	14.8	82.2	83499	9	AC026315	AC026315 Homo sapi
278	14.8	82.2	3444	10	MMU487960	AX487521 Mus muscu	C 351	14.8	82.2	83499	9	AC026315	AC026315 Homo sapi
279	14.8	82.2	3511	6	CQ585996	CQ585996 Sequence	C 352	14.8	82.2	83499	9	AC026315	AC026315 Homo sapi
280	14.8	82.2	3511	6	AX777374	AX777374 Sequence	C 353	14.8	82.2	83499	9	AC026315	AC026315 Homo sapi
281	14.8	82.2	3511	6	AX777374	AX777374 Sequence	C 354	14.8	82.2	83499	9	AC026315	AC026315 Homo sapi
282	14.8	82.2	3521	6	AX051763	AX051763 Drosophi	C 355	14.8	82.2	86358	2	AP003739	AP003739 Rattus no
C 283	14.8	82.2	3693	6	CQ600860	CQ600860 Sequence	C 356	14.8	82.2	86710	8	ATF23E12	ATF23E12 Homo sapi
C 284	14.8	82.2	3693	6	CQ600860	CQ600860 Sequence	C 357	14.8	82.2	92034	2	AC098131	AC098131 Homo sapi
285	14.8	82.2	3864	3	BT010268	BT010268 Drosophi	C 358	14.8	82.2	92187	8	AP003770	AP003770 Homo sapi
286	14.8	82.2	4624	6	CQ603334	CQ603334 Sequence	C 359	14.8	82.2	93629	2	AC024685	AC024685 Homo sapi
287	14.8	82.2	4624	6	AF444780	AF444780 Anopheles	C 360	14.8	82.2	97857	8	AP004794	AP004794 Homo sapi
288	14.8	82.2	5512	3	CQ595820	CQ595820 Sequence	C 361	14.8	82.2	99220	8	AC014958	AC014958 Homo sapi
C 289	14.8	82.2	5631	6	CQ595820	CQ595820 Sequence	C 362	14.8	82.2	99220	8	AC014958	AC014958 Homo sapi
C 290	14.8	82.2	5824	6	CQ585995	CQ585995 Sequence	C 363	14.8	82.2	100000	9	AP000017	AP000017 Homo sapi
291	14.8	82.2	6367	5	AC125231	AC125231 Homo sapi	C 364	14.8	82.2	100000	9	AP000017	AP000017 Homo sapi
292	14.8	82.2	6368	9	AC125231	AC125231 Homo sapi	C 365	14.8	82.2	100725	9	AP000015	AP000015 Homo sapi
C 293	14.8	82.2	6636	6	AX344615	AX344615 Sequence	C 366	14.8	82.2	102543	5	HS221C9	HS221C9 Homo sapi
C 294	14.8	82.2	6711	8	AY170127	AY170127 Arabidops	C 367	14.8	82.2	103039	8	AC091292	AC091292 Homo sapi
C 295	14.8	82.2	7146	6	BACTEUBAC	CQ589907 Sequence	C 368	14.8	82.2	103475	8	AP003838	AP003838 Oryza sat
296	14.8	82.2	7442	1	BACTEUBAC	CQ589907 Sequence	C 369	14.8	82.2	104470	8	AP004137	AP004137 Homo sapi
C 297	14.8	82.2	8043	6	CQ573344	CQ573344 Sequence	C 370	14.8	82.2	106159	8	AC124041	AC124041 Homo sapi
C 298	14.8	82.2	8202	6	CQ583253	CQ583253 Sequence	C 371	14.8	82.2	106159	8	ATT4C21	ATT4C21 Homo sapi
C 299	14.8	82.2	8731	6	CQ586520	CQ586520 Sequence	C 372	14.8	82.2	107972	2	AP003839	AP003839 Oryza sat
300	14.8	82.2	9404	1	AE001797	AE001797 Thermocog	C 373	14.8	82.2	107972	2	AP003839	AP003839 Oryza sat
301	14.8	82.2	9712	2	AC020107	AC020107 Drosophi	C 374	14.8	82.2	110000	1	AE000516	AE000516 Homo sapi
C 302	14.8	82.2	10029	2	AE012811	AE012811 Chlorobiu	C 375	14.8	82.2	110000	1	AE016822	AE016822 Homo sapi
C 303	14.8	82.2	10419	2	AC002371	AC002371 Homo sapi	C 376	14.8	82.2	110000	2	AC098564	AC098564 Homo sapi
C 304	14.8	82.2	10443	1	AE013914	AE013914 Yersinia	C 377	14.8	82.2	110000	2	AC098564	AC098564 Homo sapi
C 305	14.8	82.2	10615	1	AE012316	AE012316 Xanthomon	C 378	14.8	82.2	110000	2	AP006496	AP006496 Homo sapi
306	14.8	82.2	12326	5	AY279214	AY279214 Rivulius m	C 379	14.8	82.2	110000	2	AP006496	AP006496 Homo sapi
307	14.8	82.2	14050	1	AE000731	AE000731 Aquifex a	C 380	14.8	82.2	110000	2	EX890596	EX890596 Homo sapi
C 308	14.8	82.2	14350	1	AE013856	AE013856 Yersinia	C 381	14.8	82.2	110000	2	EX890596	EX890596 Homo sapi
C 309	14.8	82.2	15437	1	AF002133	AF002133 Mycobacte	C 382	14.8	82.2	110000	2	CEY105E8	CEY105E8 Homo sapi
C 310	14.8	82.2	15684	1	AB002150	AB002150 Bacillus	C 383	14.8	82.2	110000	2	LMFCHR15_6	Continuation (7 of
C 311	14.8	82.2	15683	1	AE001748	AE001748 Thermocog	C 384	14.8	82.2	110000	2	LMFCHR32_00	Continuation (7 of

385	14.8	82.2	110000	2	LMFLCHR36.01	Continuation (2 of	458	14.8	82.2	157381	2	AC129818	AC129818 Rattus no
386	14.8	82.2	110000	2	LMFLCHR36.06	Continuation (7 of	459	14.8	82.2	157442	2	AC129829	AC129829 Homo sapi
387	14.8	82.2	110000	2	LMFLCHR36.31	Continuation (32 o	460	14.8	82.2	157960	2	AC129838	AC129838 Mus muscu
388	14.8	82.2	110000	2	CR380954_06	Continuation (7 of	461	14.8	82.2	158241	2	AC021804	AC021804 Homo sapi
389	14.8	82.2	110000	8	CR382132_22	Continuation (22 o	462	14.8	82.2	158296	2	BX530098	BX530098 Danto rer
390	14.8	82.2	110000	8	CR382132_35	Continuation (36 o	463	14.8	82.2	158971	2	AP004882	AP004882 Oryza sat
391	14.8	82.2	110000	8	AC145127_15	Continuation (16 o	464	14.8	82.2	159148	8	AP005801	AP005801 Oryza sat
392	14.8	82.2	112155	9	CR378661	Continuation (16 o	465	14.8	82.2	161144	9	AC074142	AC074142 Homo sapi
393	14.8	82.2	113176	8	AC092145	AC092145 Homo sapi	466	14.8	82.2	161273	9	AL160171	AL160171 Homo sapi
394	14.8	82.2	113382	2	AC149886	AC149886 Xenopus t	467	14.8	82.2	161547	9	ALP001929	ALP001929 Homo sapi
395	14.8	82.2	113388	2	AP000590	AP000590 Homo sapi	468	14.8	82.2	161902	10	CR450715	CR450715 Danto rer
396	14.8	82.2	117555	8	AC104285	AC104285 Oryza sat	469	14.8	82.2	161905	2	AL929034	AL929034 Mouse DNA
397	14.8	82.2	118192	8	AP004120	AP004120 Oryza sat	470	14.8	82.2	162017	9	AC093522	AC093522 Homo sapi
398	14.8	82.2	119388	2	CNS09688	AP004120 Oryza sat	471	14.8	82.2	162046	9	AC093522	AC093522 Homo sapi
399	14.8	82.2	119388	8	CNS09688	AP004120 Oryza sat	472	14.8	82.2	162449	2	AP004377	AP004377 Oryza sat
400	14.8	82.2	120391	2	CNS0808C8	AL131876 Oryza sat	473	14.8	82.2	163459	2	AC092320	AC092320 Homo sapi
401	14.8	82.2	120787	8	AC000292	AC000292 Arabidops	474	14.8	82.2	165378	2	CR381620	CR381620 Homo sapi
402	14.8	82.2	121521	2	BX572074	BX572074 Danto rer	475	14.8	82.2	165652	2	AC068010	AC068010 Homo sapi
403	14.8	82.2	123384	2	BX649279	BX649279 Zebrafish	476	14.8	82.2	165931	2	AC124032	AC124032 Rattus no
404	14.8	82.2	124918	10	AC140843	AC140843 Mus muscu	477	14.8	82.2	166046	2	AC021808	AC021808 Homo sapi
405	14.8	82.2	126062	8	AP006135	AP006135 Oryza sat	478	14.8	82.2	167292	9	AL157713	AL157713 Human DNA
406	14.8	82.2	127313	2	AC078977	AC078977 Oryza sat	479	14.8	82.2	167830	10	AL928678	AL928678 Mouse DNA
407	14.8	82.2	127341	2	AC099527	AC099527 Felis cat	480	14.8	82.2	167830	9	AC016549	AC016549 Homo sapi
408	14.8	82.2	130049	8	NC1785	AL513467 Neurospor	481	14.8	82.2	167871	5	BX649473	BX649473 Zebrafish
409	14.8	82.2	131182	10	AC132255	AC132255 Mus muscu	482	14.8	82.2	168583	3	AL1354920	AL1354920 Drosophill
410	14.8	82.2	131184	10	AL844513	AL844513 Mouse DNA	483	14.8	82.2	168583	3	AL1354920	AL1354920 Human DNA
411	14.8	82.2	132060	2	AC138555	AC138555 Carolinia	484	14.8	82.2	169123	2	CR376728	CR376728 Danto rer
412	14.8	82.2	132729	8	AP005849	AP005849 Oryza sat	485	14.8	82.2	169903	2	AC069164	AC069164 Homo sapi
413	14.8	82.2	133309	2	AL355004	AL355004 Human DNA	486	14.8	82.2	170129	2	AC023712	AC023712 Drosophill
414	14.8	82.2	134378	2	AC079003	AC079003 Homo sapi	487	14.8	82.2	170761	2	AC129115	AC129115 Rattus no
415	14.8	82.2	135226	2	AC148970	AC148970 Strongylio	488	14.8	82.2	170896	10	AC140199	AC140199 Mus muscu
416	14.8	82.2	135226	8	AC105320	AC105320 Oryza sat	489	14.8	82.2	170956	2	AC127007	AC127007 Rattus no
417	14.8	82.2	137659	8	AL139232	AL139232 Human DNA	490	14.8	82.2	171032	2	AC091761	AC091761 Felis cat
418	14.8	82.2	139630	10	AC139345	AC139345 Mus muscu	491	14.8	82.2	171653	2	AC012337	AC012337 Homo sapi
419	14.8	82.2	140282	8	AC108504	AC108504 Oryza sat	492	14.8	82.2	172153	2	AC146287	AC146287 Xenopus t
420	14.8	82.2	140948	2	AC091795	AC091795 Felis cat	493	14.8	82.2	172437	2	AC022770	AC022770 Homo sapi
421	14.8	82.2	141384	2	AL358194	AL358194 Homo sapi	494	14.8	82.2	173165	3	AC099005	AC099005 Drosophill
422	14.8	82.2	141926	2	AC140687	AC140687 Rattus no	495	14.8	82.2	173192	8	BX927386	BX927386 Danto rer
423	14.8	82.2	143200	2	AC026119	AC026119 Human DNA	496	14.8	82.2	173522	2	OSJN00060	OSJN00060 Oryza sat
424	14.8	82.2	143200	2	AC026119	AC026119 Homo sapi	497	14.8	82.2	174028	8	AC087363	AC087363 Homo sapi
425	14.8	82.2	143200	8	OSJN00060	AC026119 Homo sapi	498	14.8	82.2	174314	3	AC011066	AC011066 Drosophill
426	14.8	82.2	143200	8	OSJN00060	AC026119 Homo sapi	499	14.8	82.2	174735	9	AC104449	AC104449 Homo sapi
427	14.8	82.2	143200	8	OSJN00060	AC026119 Homo sapi	500	14.8	82.2	176329	3	AC006402	AC006402 Drosophill
428	14.8	82.2	143200	8	OSJN00060	AC026119 Homo sapi	501	14.8	82.2	176802	2	AC012383	AC012383 Homo sapi
429	14.8	82.2	143200	8	OSJN00060	AC026119 Homo sapi	502	14.8	82.2	176992	2	AC111177	AC111177 Mus muscu
430	14.8	82.2	143200	8	OSJN00060	AC026119 Homo sapi	503	14.8	82.2	176992	2	AC111177	AC111177 Mus muscu
431	14.8	82.2	143200	8	OSJN00060	AC026119 Homo sapi	504	14.8	82.2	176992	2	AC111177	AC111177 Mus muscu
432	14.8	82.2	143200	8	OSJN00060	AC026119 Homo sapi	505	14.8	82.2	176992	2	AC111177	AC111177 Mus muscu
433	14.8	82.2	143200	8	OSJN00060	AC026119 Homo sapi	506	14.8	82.2	176992	2	AC111177	AC111177 Mus muscu
434	14.8	82.2	143200	8	OSJN00060	AC026119 Homo sapi	507	14.8	82.2	176992	10	AC099640	AC099640 Mus muscu
435	14.8	82.2	143200	8	OSJN00060	AC026119 Homo sapi	508	14.8	82.2	176992	9	AC016705	AC016705 Homo sapi
436	14.8	82.2	143200	8	OSJN00060	AC026119 Homo sapi	509	14.8	82.2	176992	9	AC016705	AC016705 Homo sapi
437	14.8	82.2	143200	8	OSJN00060	AC026119 Homo sapi	510	14.8	82.2	176992	9	AC016705	AC016705 Homo sapi
438	14.8	82.2	143200	8	OSJN00060	AC026119 Homo sapi	511	14.8	82.2	176992	9	AC016705	AC016705 Homo sapi
439	14.8	82.2	143200	8	OSJN00060	AC026119 Homo sapi	512	14.8	82.2	176992	9	AC016705	AC016705 Homo sapi
440	14.8	82.2	143200	8	OSJN00060	AC026119 Homo sapi	513	14.8	82.2	176992	8	AP003827	AP003827 Oryza sat
441	14.8	82.2	143200	8	OSJN00060	AC026119 Homo sapi	514	14.8	82.2	176992	2	AC093338	AC093338 Mus muscu
442	14.8	82.2	143200	8	OSJN00060	AC026119 Homo sapi	515	14.8	82.2	176992	9	AC008280	AC008280 Homo sapi
443	14.8	82.2	143200	8	OSJN00060	AC026119 Homo sapi	516	14.8	82.2	176992	9	AC021436	AC021436 Homo sapi
444	14.8	82.2	143200	8	OSJN00060	AC026119 Homo sapi	517	14.8	82.2	176992	9	AC007521	AC007521 Drosophill
445	14.8	82.2	143200	8	OSJN00060	AC026119 Homo sapi	518	14.8	82.2	176992	9	AC007521	AC007521 Drosophill
446	14.8	82.2	143200	8	OSJN00060	AC026119 Homo sapi	519	14.8	82.2	176992	9	AC007521	AC007521 Drosophill
447	14.8	82.2	143200	8	OSJN00060	AC026119 Homo sapi	520	14.8	82.2	176992	2	BX548162	BX548162 Danto rer
448	14.8	82.2	143200	8	OSJN00060	AC026119 Homo sapi	521	14.8	82.2	176992	2	AC147355	AC147355 Xenopus t
449	14.8	82.2	143200	8	OSJN00060	AC026119 Homo sapi	522	14.8	82.2	176992	2	AC147169	AC147169 Homo sapi
450	14.8	82.2	143200	8	OSJN00060	AC026119 Homo sapi	523	14.8	82.2	176992	2	AC090442	AC090442 Homo sapi
451	14.8	82.2	143200	8	OSJN00060	AC026119 Homo sapi	524	14.8	82.2	176992	3	AL135872	AL135872 Homo sapi
452	14.8	82.2	143200	8	OSJN00060	AC026119 Homo sapi	525	14.8	82.2	176992	3	AC103548	AC103548 Drosophill
453	14.8	82.2	143200	8	OSJN00060	AC026119 Homo sapi	526	14.8	82.2	176992	2	AC093198	AC093198 Drosophill
454	14.8	82.2	143200	8	OSJN00060	AC026119 Homo sapi	527	14.8	82.2	176992	3	AC118882	AC118882 Rattus no
455	14.8	82.2	143200	8	OSJN00060	AC026119 Homo sapi	528	14.8	82.2	176992	5	AC090471	AC090471 Homo sapi
456	14.8	82.2	143200	8	OSJN00060	AC026119 Homo sapi	529	14.8	82.2	176992	5	BX005201	BX005201 Zebrafish
457	14.8	82.2	143200	8	OSJN00060	AC026119 Homo sapi	530	14.8	82.2	176992	2	AC104910	AC104910 Mus muscu

C 531	14.8	82.2	185330	2	AC147410	Homo sapi	604	14.8	82.2	227370	2	AC114867	AC114867	Rattus no
C 532	14.8	82.2	185790	3	AC091221	Drosophila	C 605	14.8	82.2	227854	2	AC122584	AC122584	Rattus no
C 533	14.8	82.2	186104	9	AP005718	Homo sapi	C 606	14.8	82.2	228231	2	AC096340	AC096340	Rattus no
C 534	14.8	82.2	186341	9	AC023351	Homo sapi	C 607	14.8	82.2	228978	2	AC094027	AC094027	Rattus no
C 535	14.8	82.2	186517	2	AC116386	Mus muscu	C 608	14.8	82.2	229076	2	AC107142	AC107142	Rattus no
C 536	14.8	82.2	186541	2	AC017057	Homo sapi	C 609	14.8	82.2	229247	10	AC098888	AC098888	Mus muscu
C 537	14.8	82.2	187555	3	AC010037	Drosophila	C 610	14.8	82.2	229726	2	AC095735	AC095735	Rattus no
C 538	14.8	82.2	188489	3	AC012098	Drosophila	C 611	14.8	82.2	229872	10	AL591514	AL591514	Mouse DNA
C 539	14.8	82.2	188716	5	AL807739	Zebrafish	C 612	14.8	82.2	230926	2	AC114457	AC114457	Rattus no
C 540	14.8	82.2	190853	5	BX005041	Zebrafish	C 613	14.8	82.2	231254	5	BX784023	BX784023	Zebrafish
C 541	14.8	82.2	191046	10	AL607146	Mouse DNA	C 614	14.8	82.2	231439	2	AC111629	AC111629	Rattus no
C 542	14.8	82.2	191598	5	BX537308	Danio rer	C 615	14.8	82.2	231679	2	AC123319	AC123319	Rattus no
C 543	14.8	82.2	191785	5	BX005113	Zebrafish	C 616	14.8	82.2	232936	2	AC105642	AC105642	Rattus no
C 544	14.8	82.2	191851	2	AC149099	Saimo sal	C 617	14.8	82.2	233000	2	AC146694	AC146694	Pan trogl
C 545	14.8	82.2	193371	2	AC098398	Rattus no	C 618	14.8	82.2	233422	10	AC136675	AC136675	Mus muscu
C 546	14.8	82.2	194301	2	AP000896	Homo sapi	C 619	14.8	82.2	234431	2	AC113307	AC113307	Rattus no
C 547	14.8	82.2	195651	5	BX294114	Zebrafish	C 620	14.8	82.2	236281	2	AC112341	AC112341	Rattus no
C 548	14.8	82.2	196115	2	AC148460	Xenopus t	C 621	14.8	82.2	236783	2	AC106064	AC106064	Rattus no
C 549	14.8	82.2	196417	2	CR388495	Danio rer	C 622	14.8	82.2	239701	2	AC094922	AC094922	Rattus no
C 550	14.8	82.2	196428	2	AC148723	Saimo sal	C 623	14.8	82.2	239893	2	AC107521	AC107521	Rattus no
C 551	14.8	82.2	197424	2	AC148616	Saimo sal	C 624	14.8	82.2	239940	2	AC123752	AC123752	Mus muscu
C 552	14.8	82.2	197859	8	ATCR1V83	Alti6187 Arabidops	C 625	14.8	82.2	240483	2	AC127628	AC127628	Rattus no
C 553	14.8	82.2	197890	2	AC149092	Pan trogl	C 626	14.8	82.2	240729	2	AC111961	AC111961	Rattus no
C 554	14.8	82.2	200105	2	AP002775	Homo sapi	C 627	14.8	82.2	241178	2	AC130508	AC130508	Rattus no
C 555	14.8	82.2	201556	2	AC021757	Homo sapi	C 628	14.8	82.2	243361	2	AC094132	AC094132	Rattus no
C 556	14.8	82.2	201320	9	AC146868	Xenopus t	C 629	14.8	82.2	243846	2	AC128864	AC128864	Rattus no
C 557	14.8	82.2	201322	9	AC091194	Homo sapi	C 630	14.8	82.2	246237	3	CEY3396A	CEY3396A	Caenorhab
C 558	14.8	82.2	201528	2	AC141486	Rattus no	C 631	14.8	82.2	246852	2	AC115675	AC115675	Rattus no
C 559	14.8	82.2	202039	10	AL844182	Mouse DNA	C 632	14.8	82.2	249682	2	AC098402	AC098402	Rattus no
C 560	14.8	82.2	202328	10	AC102341	Saimo sal	C 633	14.8	82.2	250865	5	BX248410	BX248410	Zebrafish
C 561	14.8	82.2	202360	9	AC055872	Homo sapi	C 634	14.8	82.2	251301	2	AC094403	AC094403	Rattus no
C 562	14.8	82.2	202367	2	AC074394	Homo sapi	C 635	14.8	82.2	252395	2	AC130744	AC130744	Rattus no
C 563	14.8	82.2	202861	2	AC103407	Mus muscu	C 636	14.8	82.2	252985	2	AC110727	AC110727	Mus muscu
C 564	14.8	82.2	203876	2	AC010413	Homo sapi	C 637	14.8	82.2	253440	2	AC107744	AC107744	Vibrio vu
C 565	14.8	82.2	204175	2	AC128278	Rattus no	C 638	14.8	82.2	254650	1	AC096032	AC096032	Rattus no
C 566	14.8	82.2	204952	2	AC148618	Saimo sal	C 639	14.8	82.2	255447	2	AC128889	AC128889	Rattus no
C 567	14.8	82.2	205893	2	AC139999	Rattus no	C 640	14.8	82.2	259354	2	BX936422	BX936422	Danio rer
C 568	14.8	82.2	207347	9	BS000126	Pan trogl	C 641	14.8	82.2	259982	2	AL627281	AL627281	Salmoneil
C 569	14.8	82.2	208370	2	AC132123	Mus muscu	C 642	14.8	82.2	265050	1	AC126716	AC126716	Rattus no
C 570	14.8	82.2	208406	2	AC133405	Rattus no	C 643	14.8	82.2	265341	3	AE003634	AE003634	Drosophila
C 571	14.8	82.2	209405	2	BX501887	Danio rer	C 644	14.8	82.2	266132	2	AC099138	AC099138	Rattus no
C 572	14.8	82.2	209600	2	AC123601	Mus muscu	C 645	14.8	82.2	266374	2	AC118769	AC118769	Rattus no
C 573	14.8	82.2	209873	1	AL141446	Yersinia	C 646	14.8	82.2	271178	3	CEY1058A	CEY1058A	Caenorhab
C 574	14.8	82.2	210050	2	AC010732	Homo sapi	C 647	14.8	82.2	276077	3	AC103183	AC103183	Rattus no
C 575	14.8	82.2	210164	2	AP001271	Homo sapi	C 648	14.8	82.2	280895	2	AE003556	AE003556	Drosophila
C 576	14.8	82.2	210402	9	BX004964	Zebrafish	C 649	14.8	82.2	281917	3	AE003485	AE003485	Drosophila
C 577	14.8	82.2	211018	5	AC118425	Rattus no	C 650	14.8	82.2	285806	2	CH391920	CH391920	Danio rer
C 578	14.8	82.2	211018	5	BX004964	Zebrafish	C 651	14.8	82.2	286550	1	SMES91785	SMES91785	Salmoneil
C 579	14.8	82.2	211411	1	BS000001	Bacillus su	C 652	14.8	82.2	286550	2	AC116074	AC116074	Rattus no
C 580	14.8	82.2	215813	2	AC126651	Mus muscu	C 653	14.8	82.2	290285	2	AE011733	AE011733	Yersinia
C 581	14.8	82.2	215813	2	AC126651	Mus muscu	C 654	14.8	82.2	290874	1	AE013720	AE013720	Rattus no
C 582	14.8	82.2	216120	2	AC111957	Rattus no	C 655	14.8	82.2	291370	2	AC115411	AC115411	Rattus no
C 583	14.8	82.2	216531	2	AC111957	Rattus no	C 656	14.8	82.2	291601	2	AE017137	AE017137	Yersinia
C 584	14.8	82.2	217158	2	AC120575	Rattus no	C 657	14.8	82.2	291817	1	AE011857	AE011857	Rattus no
C 585	14.8	82.2	217205	8	AP004168	Oryza sat	C 658	14.8	82.2	296282	2	AE017052	AE017052	Rattus no
C 586	14.8	82.2	218474	2	AC111854	Rattus no	C 659	14.8	82.2	300029	8	AE016847	AE016847	Salmoneil
C 587	14.8	82.2	218689	2	CR388047	Danio rer	C 660	14.8	82.2	300102	1	AE017231	AE017231	Rattus no
C 588	14.8	82.2	219558	2	AC124666	Mus muscu	C 661	14.8	82.2	301068	1	AE011231	AE011231	Rattus no
C 589	14.8	82.2	220007	2	AC148935	Pan trogl	C 662	14.8	82.2	301068	1	AE011231	AE011231	Rattus no
C 590	14.8	82.2	220275	10	AL732404	Mouse DNA	C 663	14.8	82.2	301443	1	AE011231	AE011231	Rattus no
C 591	14.8	82.2	220467	10	AL732404	Mouse DNA	C 664	14.8	82.2	303250	1	AE011231	AE011231	Rattus no
C 592	14.8	82.2	221262	2	AC130225	Rattus no	C 665	14.8	82.2	303602	1	AE003433	AE003433	Rattus no
C 593	14.8	82.2	221820	2	AC133793	Rattus no	C 666	14.8	82.2	304390	2	AC111789	AC111789	Rattus no
C 594	14.8	82.2	222469	2	AC096427	Rattus no	C 667	14.8	82.2	305109	3	AE003509	AE003509	Rattus no
C 595	14.8	82.2	222810	2	AC101562	Mus muscu	C 668	14.8	82.2	305518	2	AC006879	AC006879	Caenorhab
C 596	14.8	82.2	223066	10	AC101774	Mus muscu	C 669	14.8	82.2	308050	2	BX248345	BX248345	Yersinia
C 597	14.8	82.2	223761	2	AC127239	Mus muscu	C 670	14.8	82.2	310029	1	AE016861	AE016861	Pseudomon
C 598	14.8	82.2	223764	2	AC103166	Rattus no	C 671	14.8	82.2	310581	1	AE016861	AE016861	Pseudomon
C 599	14.8	82.2	223972	2	AC027478	Homo sapi	C 672	14.8	82.2	310967	1	AE016869	AE016869	Pseudomon
C 600	14.8	82.2	224009	2	AC131058	Mus muscu	C 673	14.8	82.2	311963	1	AE016872	AE016872	Pseudomon
C 601	14.8	82.2	224915	2	AC096902	Rattus no	C 674	14.8	82.2	314432	3	AE003418	AE003418	Drosophila
C 602	14.8	82.2	225557	2	AC130243	Rattus no	C 675	14.8	82.2	316230	3	AE012593	AE012593	Rattus no
C 603	14.8	82.2	225790	2	AC099213	Rattus no	C 676	14.8	82.2	318278	3	AE003666	AE003666	Drosophila

677	14.8	82.2	322833	2	AC131863	AC131863	Rattus no	750	14.4	80.0	1976	8	SCA304452	AJ034452	Solidago
678	14.8	82.2	326728	2	AC110313	AC110313	Rattus no	751	14.4	80.0	2012	8	AJ313607	AJ313607	Homo sapi
679	14.8	82.2	334028	1	AC116537	AC116537	Drosophila	752	14.4	80.0	2050	6	CQ727169	CQ727169	Sequence
680	14.8	82.2	335050	2	AJ141450	AJ141450	Yersinia	753	14.4	80.0	2124	8	AK120224	AK120224	Oryza sat
681	14.8	82.2	340000	9	AP001730	AP001730	Homo sapi	754	14.4	80.0	2139	8	BT012895	BT012895	Sequence
682	14.8	82.2	341023	2	AC128351	AC128351	Rattus no	755	14.4	80.0	2147	1	SFL288905	SFL288905	Sequence
683	14.8	82.2	343050	1	AL935252	AL935252	Lactobacil	756	14.4	80.0	2197	10	BC021771	BC021771	Sequence
684	14.8	82.2	348764	1	BX569689	BX569689	Synechoco	757	14.4	80.0	2214	3	AK114461	AK114461	Ciona int
685	14.8	82.2	348950	1	MLEPRTN7	MLEPRTN7	Synechoco	758	14.4	80.0	2269	6	AX747402	AX747402	Sequence
686	14.8	82.2	349122	1	BX576964	BX576964	Synechoco	759	14.4	80.0	2369	6	AK092193	AK092193	Homo sapi
687	14.8	82.2	349323	1	BX572096	BX572096	Prochloro	760	14.4	80.0	2390	6	AX174729	AX174729	Sequence
688	14.8	82.2	349563	1	BX842582	BX842582	Mycobacte	761	14.4	80.0	2413	6	AX386592	AX386592	Sequence
689	14.4	80.0	1174	6	CQ806406	CQ806406	Sequence	762	14.4	80.0	2435	1	AB031322	AB031322	Sequence
690	14.4	80.0	1174	6	CQ815967	CQ815967	Sequence	763	14.4	80.0	2537	9	AK097988	AK097988	Homo sapi
691	14.4	80.0	190	9	AB051345	AB051345	Homo sapi	764	14.4	80.0	2703	1	ECOMPB	ECOMPB	Sequence
692	14.4	80.0	293	11	G73536	G73536	Dmpas Mlsc	765	14.4	80.0	2725	1	RP4TRANKF	RP4TRANKF	Sequence
693	14.4	80.0	369	6	AX184633	AX184633	Sequence	766	14.4	80.0	2745	8	AY004212	AY004212	Sequence
694	14.4	80.0	390	6	CQ748999	CQ748999	Sequence	767	14.4	80.0	2823	8	AP053104	AP053104	Nicotiana
695	14.4	80.0	439	6	AX187497	AX187497	Sequence	768	14.4	80.0	2853	6	AX286518	AX286518	Sequence
696	14.4	80.0	468	8	AY491495	AY491495	Venturia	769	14.4	80.0	2982	8	AK118117	AK118117	Sequence
697	14.4	80.0	469	8	CPA431453	CPA431453	Canis fam	770	14.4	80.0	3001	9	BC028392	BC028392	Sequence
698	14.4	80.0	508	11	G89316	G89316	S208P613RD	771	14.4	80.0	3020	9	BC030508	BC030508	Sequence
699	14.4	80.0	510	6	AX186296	AX186296	Sequence	772	14.4	80.0	3120	1	BACAKP	BACAKP	Sequence
700	14.4	80.0	556	6	CQ725781	CQ725781	Sequence	773	14.4	80.0	3148	2	AC019363	AC019363	Sequence
701	14.4	80.0	591	6	BD187016	BD187016	ES cell a	774	14.4	80.0	3296	6	BD175474	BD175474	Sequence
702	14.4	80.0	591	6	BD187024	BD187024	ES cell a	775	14.4	80.0	3296	6	AR410854	AR410854	Sequence
703	14.4	80.0	591	10	MPH34MPA	MPH34MPA	M. musculus	776	14.4	80.0	3296	6	AR439218	AR439218	Sequence
704	14.4	80.0	593	10	AF450349	AF450349	M. musculus	777	14.4	80.0	3296	6	AR473238	AR473238	Sequence
705	14.4	80.0	595	11	G61734	G61734	SHGC-87084	778	14.4	80.0	3296	6	AX464236	AX464236	Sequence
706	14.4	80.0	642	12	AY657582	AY657582	Synthetic	779	14.4	80.0	3296	6	AX697719	AX697719	Sequence
707	14.4	80.0	648	6	AX617802	AX617802	Sequence	780	14.4	80.0	3296	6	BD075623	BD075623	Sequence
708	14.4	80.0	684	6	AX412600	AX412600	Sequence	781	14.4	80.0	3296	6	BD172483	BD172483	Sequence
709	14.4	80.0	684	6	AX506268	AX506268	Sequence	782	14.4	80.0	3296	6	BD172802	BD172802	Sequence
710	14.4	80.0	684	6	AX651867	AX651867	Sequence	783	14.4	80.0	3296	6	BD173121	BD173121	Sequence
711	14.4	80.0	684	6	BT014828	BT014828	Arabidops	784	14.4	80.0	3296	6	BD173440	BD173440	Sequence
712	14.4	80.0	706	1	AY269009	AY269009	Lactobacil	785	14.4	80.0	3296	9	AY358294	AY358294	Sequence
713	14.4	80.0	740	6	BD220574	BD220574	Human gen	786	14.4	80.0	3389	8	BT002044	BT002044	Arabidops
714	14.4	80.0	780	8	CNS01993	CNS01993	Sequence	787	14.4	80.0	3889	8	PSY250467	PSY250467	Pinus syl
715	14.4	80.0	783	1	AF164439	AF164439	Bcl-2	788	14.4	80.0	4074	6	CQ806432	CQ806432	Sequence
716	14.4	80.0	783	6	AR352968	AR352968	Sequence	789	14.4	80.0	4211	8	KLCPL1	KLCPL1	Sequence
717	14.4	80.0	805	5	BX933255	BX933255	Gallus ga	790	14.4	80.0	4338	6	CO815989	CO815989	Sequence
718	14.4	80.0	808	6	AX188261	AX188261	Sequence	791	14.4	80.0	4488	6	AR354110	AR354110	Sequence
719	14.4	80.0	812	8	BT012597	BT012597	Arabidops	792	14.4	80.0	4545	2	AC017672	AC017672	Sequence
720	14.4	80.0	819	6	AX620256	AX620256	Sequence	793	14.4	80.0	4940	6	A92451	A92451	Sequence
721	14.4	80.0	825	5	AX358873	AX358873	Canis fam	794	14.4	80.0	4940	12	CVE131294	CVE131294	Sequence
722	14.4	80.0	870	5	BX929405	BX929405	Gallus ga	795	14.4	80.0	5239	8	OSAS35058	OSAS35058	Sequence
723	14.4	80.0	926	11	CNS0610W	CNS0610W	Sequence	796	14.4	80.0	5239	2	AC017408	AC017408	Sequence
724	14.4	80.0	951	8	AK104036	AK104036	Oryza sat	797	14.4	80.0	6570	6	CO574688	CO574688	Sequence
725	14.4	80.0	951	8	AK104036	AK104036	Oryza sat	798	14.4	80.0	6778	6	CO598244	CO598244	Sequence
726	14.4	80.0	1068	3	AY008134	AY008134	Caenorhab	799	14.4	80.0	7401	6	CO598244	CO598244	Sequence
727	14.4	80.0	1109	8	AY008134	AY008134	Caenorhab	800	14.4	80.0	7597	2	AC017434	AC017434	Sequence
728	14.4	80.0	1115	8	AK060975	AK060975	Oryza sat	801	14.4	80.0	7930	9	HSB808084	HSB808084	Sequence
729	14.4	80.0	1159	6	BD269309	BD269309	33 human	802	14.4	80.0	10431	1	AE002318	AE002318	Homo sapi
730	14.4	80.0	1200	6	BD269287	BD269287	33 human	803	14.4	80.0	11037	1	AE006359	AE006359	Sequence
731	14.4	80.0	1397	5	BX935076	BX935076	Gallus ga	804	14.4	80.0	11067	2	AC014725	AC014725	Sequence
732	14.4	80.0	1413	6	AX838634	AX838634	Sequence	805	14.4	80.0	1167	2	AE004536	AE004536	Sequence
733	14.4	80.0	1462	6	AR354653	AR354653	Sequence	806	14.4	80.0	11905	1	AE006347	AE006347	Sequence
734	14.4	80.0	1474	6	AF056283	AF056283	Nicotiana	807	14.4	80.0	12464	3	CEC44E1	CEC44E1	Sequence
735	14.4	80.0	1521	3	AP495633	AP495633	Caenantho	808	14.4	80.0	12588	3	CEC44E1	CEC44E1	Sequence
736	14.4	80.0	1521	3	AP495633	AP495633	Caenantho	809	14.4	80.0	14456	1	AY513488	AY513488	Sequence
737	14.4	80.0	1521	3	AP495633	AP495633	Caenantho	810	14.4	80.0	14787	1	AY513488	AY513488	Sequence
738	14.4	80.0	1521	3	AP495633	AP495633	Caenantho	811	14.4	80.0	14810	2	AC014701	AC014701	Sequence
739	14.4	80.0	1521	3	AP495633	AP495633	Caenantho	812	14.4	80.0	15702	14	AJ508938	AJ508938	Sequence
740	14.4	80.0	1521	3	AP495633	AP495633	Caenantho	813	14.4	80.0	16341	1	AY048850	AY048850	Sequence
741	14.4	80.0	1521	3	AP495633	AP495633	Caenantho	814	14.4	80.0	16341	1	AY048850	AY048850	Sequence
742	14.4	80.0	1521	3	AP495633	AP495633	Caenantho	815	14.4	80.0	19041	12	AF327712	AF327712	Cloning v
743	14.4	80.0	1521	3	AP495633	AP495633	Caenantho	816	14.4	80.0	20035	2	SPNEU1915	SPNEU1915	Sequence
744	14.4	80.0	1521	3	AP495633	AP495633	Caenantho	817	14.4	80.0	21159	2	AC020461	AC020461	Sequence
745	14.4	80.0	1521	3	AP495633	AP495633	Caenantho	818	14.4	80.0	21583	1	AE008838	AE008838	Salmonell
746	14.4	80.0	1521	3	AP495633	AP495633	Caenantho	819	14.4	80.0	24091	9	AC008010	AC008010	Homo sapi
747	14.4	80.0	1521	3	AP495633	AP495633	Caenantho	820	14.4	80.0	27608	12	AF184578	AF184578	Homo sapi
748	14.4	80.0	1521	3	AP495633	AP495633	Caenantho	821	14.4	80.0	27719	9	AC012623	AC012623	Homo sapi
749	14.4	80.0	1521	3	AP495633	AP495633	Caenantho	822	14.4	80.0	28582	9	AC104822	AC104822	Homo sapi

C 823	14.4	80.0	26709.12	AY128787 Cloning v
C 824	14.4	80.0	26935.10	AC120382 Mms muscu
C 825	14.4	80.0	30660.10	CR548624 Mouse DNA
C 826	14.4	80.0	31776	UJ9855 Caenorhabd
C 827	14.4	80.0	36120	UJ9997 Caenorhabd
C 828	14.4	80.0	36396	AF039713 Caenorhab
C 829	14.4	80.0	36697	AB027255 Cloning v
C 830	14.4	80.0	37013	SPAC3A12 S.pombe chr
C 831	14.4	80.0	37248	CEP46F2 Z69903 Caenorhabd
C 832	14.4	80.0	38188	AY120852 AC120852 Symechnoc
C 833	14.4	80.0	39132	CBRG45T03 AC0084637 Caenorhab
C 834	14.4	80.0	39283	CEI704A11 ZB3123 Caenorhabd
C 835	14.4	80.0	39355	CBRG05U05 AC0084488 Caenorhab
C 836	14.4	80.0	39456	AC0792983 AC0792983 Arabidops
C 837	14.4	80.0	39512	AY4480105 Ipomoea t
C 838	14.4	80.0	39502	AB027254 Cloning v
C 839	14.4	80.0	40172	AF022968 Caenorhab
C 840	14.4	80.0	41293	Z62791 Caenorhabd
C 841	14.4	80.0	43657	AC0097506 AC0097506 Homo sapi
C 842	14.4	80.0	43883	AC0043396 AC0043396 pseudomon
C 843	14.4	80.0	44663	U49830 Caenorhabd
C 844	14.4	80.0	46622	AC118624 AC118624 Mms muscu
C 845	14.4	80.0	48128	AX059513 Sequence
C 846	14.4	80.0	49393	AC101083 Mms muscu
C 847	14.4	80.0	52593	CEY45E108 AC1021487 Caenorhab
C 848	14.4	80.0	55234	AC0933023 AC0933023 Homo sapi
C 849	14.4	80.0	57245	AB010069 AB010069 Arabidops
C 850	14.4	80.0	57367	AC110055 AC110055 Homo sapi
C 851	14.4	80.0	58479	ALV77281 Mouse DNA
C 852	14.4	80.0	58672	AC006524 Homo sapi
C 853	14.4	80.0	58847	AC024869 Caenorhab
C 854	14.4	80.0	59264	AC018278 Drosophil
C 855	14.4	80.0	59766	AC100128 Mms muscu
C 856	14.4	80.0	60099	B1ACGEXN L27758 Birmingham
C 857	14.4	80.0	61182	AC100071 AC100071 Mms muscu
C 858	14.4	80.0	61984	AL353557 Human DNA
C 859	14.4	80.0	62276	AL929118 Mouse DNA
C 860	14.4	80.0	65838	AC115758 Mus muscu
C 861	14.4	80.0	66223	AC137858 Mus muscu
C 862	14.4	80.0	66596	AC106801 AC106801 Mms muscu
C 863	14.4	80.0	68304	BX000432 Mouse DNA
C 864	14.4	80.0	72221	AC090331 Homo sapi
C 865	14.4	80.0	72438	ECOTW67.3 Continuation (4 of
C 866	14.4	80.0	72533	AC013615 AC013615 Homo sapi
C 867	14.4	80.0	72631	AL929299 ZedraFish
C 868	14.4	80.0	72666	HSDJ581F3 AL105805 Human DNA
C 869	14.4	80.0	74201	AC0084252 AF084252 Homo sapi
C 870	14.4	80.0	74406	AC139230 AC139230 Arabidops
C 871	14.4	80.0	74706	AC100601 AC100601 Mms muscu
C 872	14.4	80.0	74832	AB025604 Arabidops
C 873	14.4	80.0	75375	AC0021969 AC0021969 Homo sapi
C 874	14.4	80.0	75376	AC104065 Homo sapi
C 875	14.4	80.0	77127	AC016804 Homo sapi
C 876	14.4	80.0	77547	AB011474 Arabidops
C 877	14.4	80.0	79899	CR547127 Mms muscu
C 878	14.4	80.0	80255	AL929550 Mouse DNA
C 879	14.4	80.0	80734	AP001537 Homo sapi
C 880	14.4	80.0	81450	AL590346 Arabidops
C 881	14.4	80.0	82896	AC004744 AC004744 Homo sapi
C 882	14.4	80.0	85579	ATTT24H18 AL353013 Arabidops
C 883	14.4	80.0	90020	AL354758 Human DNA
C 884	14.4	80.0	92625	AL354758 Human DNA
C 885	14.4	80.0	92626	AL149155 Xenopus t
C 886	14.4	80.0	95266	AL022603 Arabidops
C 887	14.4	80.0	95267	AL149154 Xentopus t
C 888	14.4	80.0	95845	AF411807 Lycopersis
C 889	14.4	80.0	96057	AC0066217 Arabidops
C 890	14.4	80.0	96624	BK004802 Zebrafish
C 891	14.4	80.0	96925	AC0017831 Drosophil
C 892	14.4	80.0	97065	ABD00

C 969	14.4	80.0	131152	10	AL929068	AL929068 Mouse DNA
C 970	14.4	80.0	131441	8	CNS08BCDA	AL954871 Oryza sat
C 971	14.4	80.0	131692	8	AC006423	AC006423 Arabidops
C 972	14.4	80.0	132728	9	AC024121	AC024121 Homo sapi
C 973	14.4	80.0	133568	9	HS16915	233015 Human DNA s
C 974	14.4	80.0	133615	2	AC146749	AC146749 Medicago
C 975	14.4	80.0	134355	8	AC124952	AC124952 Medicago
C 976	14.4	80.0	134532	9	AP006442	AP006442 Homo sapi
C 977	14.4	80.0	134712	2	AF003891	AF003891 Oryza sat
C 978	14.4	80.0	134717	10	AL626805	AL626805 Mouse DNA
C 979	14.4	80.0	134741	10	AL683819	AL683819 Mouse DNA
C 980	14.4	80.0	134938	2	AC137759	AC137759 Oryza sat
C 981	14.4	80.0	135105	8	CNS08CAB	AL831795 Oryza sat
C 982	14.4	80.0	135660	2	AC103927	AC103927 Mus muscu
C 983	14.4	80.0	135820	2	AL513497	AL513497 Human DNA
C 984	14.4	80.0	136657	10	AC124572	AC124572 Mus muscu
C 985	14.4	80.0	137122	10	AC123037	AC123037 Mus muscu
C 986	14.4	80.0	137674	2	AC117697	AC117697 Mus muscu
C 987	14.4	80.0	138020	9	AC108043	AC108043 Homo sapi
C 988	14.4	80.0	138231	9	AC093558	AC093558 Homo sapi
C 989	14.4	80.0	138264	9	AC008722	AC008722 Homo sapi
C 990	14.4	80.0	138633	2	AL356360	AL356360 Homo sapi
C 991	14.4	80.0	139084	5	AL935332	AL935332 Zabrafish
C 992	14.4	80.0	139408	2	AC078565	AC078565 Mus muscu
C 993	14.4	80.0	140750	2	AC131239	AC131239 Medicago
C 994	14.4	80.0	140944	10	AL450331	AL450331 Mouse DNA
C 995	14.4	80.0	141458	2	AC142044	AC142044 Rattus no
C 996	14.4	80.0	142273	9	AL391994	AL391994 Human DNA
C 997	14.4	80.0	142475	8	AP003627	AP003627 Oryza sat
C 998	14.4	80.0	143206	2	AC141776	AC141776 Apis mell
C 999	14.4	80.0	143575	2	AC092370	AC092370 Homo sapi
C1000	14.4	80.0	143765	8	AP004623	AP004623 Oryza sat

ALIGNMENTS

RESULT 1
CR382126 06/c

WPCOMMENT
Sequence split into 26 fragments LOCUS CR382126 Accession CR382126

Fragment Name	Begin	End
CR382126_00	1	110000
CR382126_01	100001	210000
CR382126_02	200001	310000
CR382126_03	300001	410000
CR382126_04	400001	510000
CR382126_05	500001	610000
CR382126_06	600001	710000
CR382126_07	700001	810000
CR382126_08	800001	900001
CR382126_09	900001	1010000
CR382126_10	1000001	1100001
CR382126_11	1100001	1210000
CR382126_12	1200001	1310000
CR382126_13	1300001	1410000
CR382126_14	1400001	1510000
CR382126_15	1500001	1610000
CR382126_16	1600001	1710000
CR382126_17	1700001	1810000
CR382126_18	1800001	1910000
CR382126_19	1900001	2010000
CR382126_20	2000001	2110000
CR382126_21	2100001	2210000
CR382126_22	2200001	2310000
CR382126_23	2300001	2410000
CR382126_24	2400001	2510000
CR382126_25	2500001	2602197

Continuation (7 of 26) of CR382126 from base 600001 (CR382126 Kluyveromyces fragilis strain
Query Match 94.4%; Score 17; DB 8; Length 110000;
Best Local Similarity 100.0%; Pred. No. 2,4e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GAGAACCAACAGATTGG 17
Db 81584 GAGAACCAACAGATTGG 81568
RESULT 2
AE006452/c 10610 bp DNA linear BCT 02-JUN-2004
LOCUS
DEFINITION
Lactococcus lactis subsp. lactis IL1403 section 214 of the
complete genome.
ACCESSION
AE006452 AE005176
VERSION
AE006452.1 GI:12725297
KEYWORDS
SOURCE
ORGANISM
Lactococcus lactis subsp. lactis IL1403
Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
Lactococcus.
REFERENCE
1 (bases 1 to 10610)
Bojotin, A., Wincker, P., Mager, S., Jaillon, O., Malarme, K.,
Weissenbach, J., Ehrlich, S.D. and Sorokin, A.
The complete genome sequence of the lactic acid bacterium
Lactococcus lactis ssp. lactis IL1403
Genome Res. 11 (5), 731-753 (2001)
MEDLINE
21235186
PUBMED
1137471
AUTHORS
2 (bases 1 to 10610)
Bojotin, A., Wincker, P., Mager, S., Jaillon, O., Malarme, K.,
Weissenbach, J., Ehrlich, S.D. and Sorokin, A.
Direct Submission
Submitted (09-JUN-2001) INRA, Genetique Microbienne, Domaine de
Vilvert, Jouy en Josas 91852, France
FEATURES
source
1..10610
/organism="Lactococcus lactis subsp. lactis IL1403"
/mol_type="genomic DNA"
/strain="IL1403"
/sub_species="lactis"
/db_xref="taxon:272623"
complement (83..610)
/gene="wreD"
/note="synonym: L98132"
complement (83..610)
/gene="wreD"
/note="EVIDENCE BY HOMOLOGO B1003.02 WreD sacculus and
peptidoglycan. INVOLVED IN FORMATION OF THE SHAPE OF THE
CELL. MAY ALSO CONTRIBUTE TO REGULATION OF FORMATION OF
PENICILLIN-BINDING PROTEINS"
/codon_start=1
/transl_table=1
/product="cell shape determining protein"
/protein_id="AAK06329.1"
/db_xref="GI:12725297"
/translation="MAKINLSKUIIIIIIIIAANSAIFAKNFKNRNSAQTQTI
NDISTFVDRVLAIPHVFQDKANEDSNIMDTYQONQSLKTOLAKSKDDNKLSTGDESE

gene
CDS
10610
/organism="Lactococcus lactis subsp. lactis IL1403"
/mol_type="genomic DNA"
/strain="IL1403"
/sub_species="lactis"
/db_xref="taxon:272623"
complement (83..610)
/gene="wreD"
/note="synonym: L98132"
complement (83..610)
/gene="wreD"
/note="EVIDENCE BY HOMOLOGO B1003.02 WreD sacculus and
peptidoglycan. INVOLVED IN FORMATION OF THE SHAPE OF THE
CELL. MAY ALSO CONTRIBUTE TO REGULATION OF FORMATION OF
PENICILLIN-BINDING PROTEINS"
/codon_start=1
/transl_table=1
/product="cell shape determining protein"
/protein_id="AAK06329.1"
/db_xref="GI:12725297"
/translation="MAKINLSKUIIIIIIIIAANSAIFAKNFKNRNSAQTQTI
NDISTFVDRVLAIPHVFQDKANEDSNIMDTYQONQSLKTOLAKSKDDNKLSTGDESE

NKEELKALKEOETLTDTQVAVANITREPPSMDTLVIDSGSKDGLTTGMITWANGV
 VGRVQVANKSKVALLSSKIDNKPVRLESQSPIGILSSYDSQOXYVVKKID
 SOGKRNQDVSFTSGITNGSGOGTPEGCLLIGVIGEKQSDGLGMMVYIKPBTILY
 DIRFVYVKSXIDQ"
 1868..2308
 /gene="yxbC"
 /note="synonym: L99869"
 1868..2308
 /gene="yxbC"
 /note="EVIDENCE BY HOMOLOGY BIO15.01 HYPOTHETICAL."
 General"
 /codon_start=1
 /transl_table=1
 /product="HYPOTHETICAL PROTEIN"
 /protein_id="AAK06330.1"
 /db_xref="GI:12725300"
 /translation="NMHI PVFENKNEKNQARYGVYSIVARENKEICLVQAPNGAYF
 LPGEIEAGDEQEKALHRELMELGATLIEFGQADEYFSSHRDQYYNDAIYIN
 TKNVFDSKPLEDPNHSIMDPETALSKLKGSHKMGVQEWLKK"
 complement(12336..3010)
 /gene="piA"
 /note="synonym: L0045"
 complement(2336..3010)
 /gene="piA"
 /note="EVIDENCE BY HOMOLOGY BIO06.10 Pentose phosphate
 pathway"
 /codon_start=1
 /transl_table=1
 /product="ribose 5-phosphate isomerase A (EC 5.3.1.6)"
 /protein_id="AAK06331.1"
 /db_xref="GI:12725301"
 /translation="MDNLKQVGIKAEFVKSQVVGIGTSTAYFVEELGRIAE
 QLEITGVTSNVSQARALGIPLASIDRVYDITVDGADEISLNGIKGGALL
 MEKIVATYKQIYVDESKLSENGISFKIPEVIYQSGQVFKFPAAGVAPYMRIN
 BENERLITDMHFIIDHISQIKPEKKAELDMGAVHGLFNNVVKVIVANGEG
 VRIIIG"
 complement(3094..3780)
 /gene="rcfB"
 /note="synonym: L0245"
 complement(3094..3780)
 /gene="rcfB"
 /note="EVIDENCE BY HOMOLOGY BIO09.01 REGULATORY FUNCTIONS.
 General. BELONGS TO THE CRP/FNR FAMILY OF TRANSCRIPTIONAL
 REGULATORS"
 /codon_start=1
 /transl_table=1
 /product="transcriptional regulator"
 /protein_id="AAK06332.1"
 /db_xref="GI:12725302"
 /translation="WSPFKMEEMANVNIYLLINTEOKKVVYTRKRSYIMYOGIES
 NYIVLKQGVAKITSLIDRGEFNIAVYADPVSILBEKONDISALFVVRVESPRA
 NYKISRDPMWVREDINLFRVYVDPFKRLANLBILOKNTNGKRGVACLSHSL
 IDDSIRKDDGLIDFTVNEIDIAFGCISTRSNVNRLHDLDEKVIDIVDKMITY
 NPOYLEYIS"
 complement(3863..5083)
 /gene="yxbD"
 /note="synonym: L101912"
 complement(3863..5083)
 /gene="yxbD"
 /note="EVIDENCE BY HOMOLOGY BIO13.08 TRANSPORT AND BINDING
 PROTEINS. Multidrug resistance, similar to quinolone
 resistance protein Nora from difference species"
 /codon_start=1
 /transl_table=1
 /product="transporter"
 /protein_id="AAK06333.1"
 /db_xref="GI:12725303"
 /translation="MKESITPTPTVNTLSIFIVIVYVLSIGTYVALQLOHST
 VVLSUSVFIYVIGALIGRWGTINIRIGKRLIYIGLIFVITRFGYITNPLF
 LRVIGQAGFGIATGATVAVPARGEGIGYALSVTLAAVGPALSLNIVS
 LGFSLIGLIGLIVITFIIFFVVKESFAEERAYALAHKPKVEYIIEKSLPISI
 IAFAGFISAILTGMSFSDIKIPLAGSLFFIMYAVLIFVSRPFTGRLLPDKGDM
 INPFTIFPAVAMLVGLAGFSPAIGFVILLIGCAVGLGAVAPFGQAIARDS

KDRIGVATTFEGFLDIGVGGPIVIGALIIIMLGNGNLGRNLYLYSAPAVIWIYI
 YLHGKQKSSSKV"
 complement(5146..5583)
 /gene="yxbE"
 /note="synonym: L103195"
 complement(5146..5583)
 /gene="yxbE"
 /note="EVIDENCE BY HOMOLOGY BIO15.02 HYPOTHETICAL.
 conserved. there are six paralogs in L.lactis"
 /codon_start=1
 /transl_table=1
 /product="conserved hypothetical protein"
 /protein_id="AAK06334.1"
 /db_xref="GI:12725304"
 /translation="MEQKISRIIVGNDGSAESLKAPDEAVRVAKRDAELILANVEK
 RSLAVNTYDTIAEHEHNGTKLLQOFARALDEGLTNRITLEYGSPRVAKATIP
 KOENVDLIVIGATGISYIERVYVGSVASYIVNATCNVILVRK"
 complement(5603..5806)
 /gene="yxbF"
 /note="synonym: L103652"
 complement(5603..5806)
 /gene="yxbF"
 /note="EVIDENCE BY HOMOLOGY BIO15.01 HYPOTHETICAL.
 General"
 /codon_start=1
 /transl_table=1
 /product="HYPOTHETICAL PROTEIN"
 /protein_id="AAK06335.1"
 /db_xref="GI:12725305"
 /translation="MKKFPKLSSEYNSIPNYVGTAKRSYQDFVNMDEALFVSN
 HHGFKTBDGKFAFERNISATG"
 complement(6066..10376)
 /gene="yxcA"
 /note="synonym: L104115"
 complement(6066..10376)
 /gene="yxcA"
 /note="EVIDENCE BY HOMOLOGY BIO15.02 HYPOTHETICAL.
 conserved. 30% identical to activator of
 (R)-2-hydroxyglutaryl-CoA dehydratase of Acidimicrococcus
 fermentans."
 /codon_start=1
 /transl_table=1
 /product="conserved hypothetical protein"
 /protein_id="AAK06336.1"
 /db_xref="GI:12725306"
 /translation="WTOIYKIGIDVSGTAVKLVFDPENFKLISRYERHSDVKAATI
 KVPDFIEYEDNLBELAITGSGMGAAEVIDVFEQVEVIASSITTEKLPOTVWIE
 LGEDAKVTFPDNGTOELRNNGTCAGGTGAFIDQVALLKVDANGVIELKKNKLYP
 IASRGVAFKTDVQPLNDEGRREDIAASIFQAVVNTTIGLNGSKIKNVAFILGSP
 LFPSELRKRIETLIDKEENIIPENPOLFVANGALYEISQKATDILHLEND
 SDSLIPOIMIKFESQALDEPRTRAKTANVKNLSEHGAALFGIDIGSTTKY
 LIDEGNLIYHNGSNGGPLESYINILKOVAKLPADVITASTVNGVEHILKAGL
 HVDGEVETMHHYKADYFNPQVDFIIDIGQDKAMRIDGALSSIQLEASSSGQ
 SFIEFASLHDVREFSQALLAEHVNIGSRCTVFNKSVKQVQREGATVDISAG
 LSYVIRKALYKIVKRPBELIKRTYVQGTIVNLSVAKVAKVQVREGATVDISAG
 MGAVGALISLIDQENANKSEILKLDLEFTYHKEVNLGALCENCKMLTVFNDG
 KFAVGNCEBAEKATYKVAKKDKKNLVADYKXKLFYHSHSKKQGTGEIGIPV
 LMYENYPLATHMLTDFGFRVLSFSDKELFEEGIETIPSDYCYPAKSHGHIAL
 IKOGIPNIFVPSVLFEOBKNAQNHNGCIYOSPEYV.LKNMIDEBEGOVNLHPFI
 NLANDGVAVVHKALTRQGISVLTETVQVAVOHGPEEMKFEEDLRLKREELMOIN
 LNDEKATYLAERPPIHLPETINHGADITIDGSHVLTEDSISHLABKSRVYNQVY
 HSRILPAANVVCNKNLLELVQNSFGGLDAVTTDOVERIMRGNKLYTLKIDBSN

Query Match 91.1% Score 16.4; DB 1; Length 10610;
 Best Local Similarity 94.4%; Pred. No. 5.6e+02;
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 1 GGAACAACAAGATCCG 18
 Db 5871 GGAACAACAAGATCCG 5854

RESULT 3

LOCUS	DEFINITION	VERSION	ACCESSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE
MLCB1788	39228 bp	DNA	linear	BCT 27-AUG-1999					
MLCB1788	Mycobacterium leprae cosmid B1788.								
AL008609	431, 6-phosphogluconate dehydrogenase; gcvH; gcvR; glcB; glycine cleavage system h protein; glycine dehydrogenase (decarboxylating); glnD; gnaB; oxalose-5'-monophosphate dehydrogenase; LEPREP; malate synthase g; oxidoreductase; psaA; phosphatidyltransferase; triphosphatase; secA; secB; secE; secF; secH; secI; secJ; secK; secL; secM; secN; secO; secP; secQ; secR; secS; secT; secU; secV; secW; secX; secY; secZ; secAA; secAB; secAC; secAD; secAE; secAF; secAG; secAH; secAI; secAJ; secAK; secAL; secAM; secAN; secAO; secAP; secAQ; secAR; secAS; secAT; secAU; secAV; secAW; secAX; secAY; secAZ; secBA; secBB; secBC; secBD; secBE; secBF; secBG; secBH; secBI; secBJ; secBK; secBL; secBM; secBN; secBO; secBP; secBQ; secBR; secBS; secBT; secBU; secBV; secBW; secBX; secBY; secBZ; secCA; secCB; secCC; secCD; secCE; secCF; secCG; secCH; secCI; secCJ; secCK; secCL; secCM; secCN; secCO; secCP; secCQ; secCR; secCS; secCT; secCU; secCV; secCW; secCX; secCY; secCZ; secDA; secDB; secDC; secDD; secDE; secDF; secDG; secDH; secDI; secDJ; secDK; secDL; secDM; secDN; secDO; secDP; secDQ; secDR; secDS; secDT; secDU; secDV; secDW; secDX; secDY; secDZ; secEA; secEB; secEC; secED; secEE; secEF; secEG; secEH; secEI; secEJ; secEK; secEL; secEM; secEN; secEO; secEP; secEQ; secER; secES; secET; secEU; secEV; secEW; secEX; secEY; secEZ; secFA; secFB; secFC; secFD; secFE; secFF; secFG; secFH; secFI; secFJ; secFK; secFL; secFM; secFN; secFO; secFP; secFQ; secFR; secFS; secFT; secFU; secFV; secFW; secFX; secFY; secFZ; secGA; secGB; secGC; secGD; secGE; secGF; secGG; secGH; secGI; secGJ; secGK; secGL; secGM; secGN; secGO; secGP; secGQ; secGR; secGS; secGT; secGU; secGV; secGW; secGX; secGY; secGZ; secHA; secHB; secHC; secHD; secHE; secHF; secHG; secHI; secHJ; secHK; secHL; secHM; secHN; secHO; secHP; secHQ; secHR; secHS; secHT; secHU; secHV; secHW; secHX; secHY; secHZ; secIA; secIB; secIC; secID; secIE; secIF; secIG; secIH; secIJ; secIK; secIL; secIM; secIN; secIO; secIP; secIQ; secIR; secIS; secIT; secIU; secIV; secIW; secIX; secIY; secIZ; secJA; secJB; secJC; secJD; secJE; secJF; secJG; secJH; secJI; secJJ; secJK; secJL; secJM; secJN; secJO; secJP; secJQ; secJR; secJS; secJT; secJU; secJV; secJW; secJX; secJY; secJZ; secKA; secKB; secKC; secKD; secKE; secKF; secKG; secKH; secKI; secKJ; secKK; secKL; secKM; secKN; secKO; secKP; secKQ; secKR; secKS; secKT; secKU; secKV; secKW; secKX; secKY; secKZ; secLA; secLB; secLC; secLD; secLE; secLF; secLG; secLH; secLI; secLJ; secLK; secLL; secLM; secLN; secLO; secLP; secLQ; secLR; secLS; secLT; secLU; secLV; secLW; secLX; secLY; secLZ; secMA; secMB; secMC; secMD; secME; secMF; secMG; secMH; secMI; secMJ; secMK; secML; secMN; secMO; secMP; secMQ; secMR; secMS; secMT; secMU; secMV; secMW; secMX; secMY; secMZ; secNA; secNB; secNC; secND; secNE; secNF; secNG; secNH; secNI; secNJ; secNK; secNL; secNM; secNO; secNP; secNQ; secNR; secNS; secNT; secNU; secNV; secNW; secNX; secNY; secNZ; secOA; secOB; secOC; secOD; secOE; secOF; secOG; secOH; secOI; secOJ; secOK; secOL; secOM; secON; secOO; secOP; secOQ; secOR; secOS; secOT; secOU; secOV; secOW; secOX; secOY; secOZ; secPA; secPB; secPC; secPD; secPE; secPF; secPG; secPH; secPI; secPJ; secPK; secPL; secPM; secPN; secPO; secPP; secPQ; secPR; secPS; secPT; secPU; secPV; secPW; secPX; secPY; secPZ; secQA; secQB; secQC; secQD; secQE; secQF; secQG; secQH; secQI; secQJ; secQK; secQL; secQM; secQN; secQO; secQP; secQQ; secQR; secQS; secQT; secQU; secQV; secQW; secQX; secQY; secQZ; secRA; secRB; secRC; secRD; secRE; secRF; secRG; secRH; secRI; secRJ; secRK; secRL; secRM; secRN; secRO; secRP; secRQ; secRR; secRS; secRT; secRU; secRV; secRW; secRX; secRY; secRZ; secSA; secSB; secSC; secSD; secSE; secSF; secSG; secSH; secSI; secSJ; secSK; secSL; secSM; secSN; secSO; secSP; secSQ; secSR; secSS; secST; secSU; secSV; secSW; secSX; secSY; secSZ; secTA; secTB; secTC; secTD; secTE; secTF; secTG; secTH; secTI; secTJ; secTK; secTL; secTM; secTN; secTO; secTP; secTQ; secTR; secTS; secTT; secTU; secTV; secTW; secTX; secTY; secTZ; secUA; secUB; secUC; secUD; secUE; secUF; secUG; secUH; secUI; secUJ; secUK; secUL; secUM; secUN; secUO; secUP; secUQ; secUR; secUS; secUT; secUU; secUV; secUW; secUX; secUY; secUZ; secVA; secVB; secVC; secVD; secVE; secVF; secVG; secVH; secVI; secVJ; secVK; secVL; secVM; secVN; secVO; secVP; secVQ; secVR; secVS; secVT; secVU; secVV; secVW; secVX; secVY; secVZ; secWA; secWB; secWC; secWD; secWE; secWF; secWG; secWH; secWI; secWJ; secWK; secWL; secWM; secWN; secWO; secWP; secWQ; secWR; secWS; secWT; secWU; secWV; secWW; secWX; secWY; secWZ; secXA; secXB; secXC; secXD; secXE; secXF; secXG; secXH; secXI; secXJ; secXK; secXL; secXM; secXN; secXO; secXP; secXQ; secXR; secXS; secXT; secXU; secXV; secXW; secXX; secXY; secXZ; secYA; secYB; secYC; secYD; secYE; secYF; secYG; secYH; secYI; secYJ; secYK; secYL; secYM; secYN; secYO; secYP; secYQ; secYR; secYS; secYT; secYU; secYV; secYW; secYX; secYY; secYZ; secZA; secZB; secZC; secZD; secZE; secZF; secZG; secZH; secZI; secZJ; secZK; secZL; secZM; secZN; secZO; secZP; secZQ; secZR; secZS; secZT; secZU; secZV; secZW; secZX; secZY; secZZ								

COMMENT

Notes: The Sanger Centre is funded to complete the sequence of M. leprae by the Heiser Program for Research in Leprosy and Tuberculosis of The New York Community Trust.

Work in Paris is supported by the Heiser Trust, the Association Française Rouli Follereau and the Groupement de Recherches et Etudes des Genomes (GIR-GEEG).

Details of M. leprae sequencing at the Sanger Centre are available on the World Wide Web

(URL: <http://www.sanger.ac.uk/Projects//>)

CDS are numbered using the following system eg MLCB33.01c. ML ('M. lepre', Cb33 (cosmid name), .01 (first CDS), c (complementary strand).

The more significant matches with motifs in the PROSIRE database are also included but some of these may be fortuitous. The length in codons is given for each CDS.

Usually the highest scoring match found by fasta-o is given for CDS which show significant similarity to other CDS in the database. The position of possible ribosome binding site sequences are given where these have been used to deduce the initiation codon. All CDS over 100 codons have been analysed. Gene prediction is based on positional base preference in codons especially where there is an increase in the observed/expected third position G + C. CAUTION: We may not have predicted the correct initiation codon. Where possible we choose an initiation codon (atg, gtg, or ttg) which is preceded by an upstream ribosome binding site sequence (optimally 5-13bp before the initiation codon). If this cannot be identified we choose the most upstream initiation codon.

IMPORTANT: This sequence MAY NOT be the entire insert of the sequenced clone. It may be shorter because we only sequence overlapping sections once, or longer, because we arrange for a small overlap between neighbouring subclones. Cosmid B1788 is overlapped at the 5' end by EM BA:MLB38COS.

FEATURES

```
/organism="Mycobacterium leprae"  
/mol_type="genomic DNA"
```

```
misc_feature
1. .2183
/db_xref="taxon:1769"
/clone="cosmid B1788"
```

gene

CDS

ma'sc

RBS

CDS

CDS

CTCS

misc

CDS


```

CDS
/pseudo
/codon_start=1
/transl_table=11
/complement(1805, .2075)
/notes="MNCB1788.07c, ABC-transporter pseudogene fragment,
len: 89 aa; highly similar to part of TR:P95156 O08061,
MTCY359.15c, modB, membrane spanning subunit of
ABC-transport system (264 aa), (60.7% identity in 89 aa
overlap)"
/pseudo
/codon_start=1
/transl_table=11
/db_xref="PSEUDO:CA15439.1"
/complement(2075, .2160)
/notes="MNCB1788.08c, ABC-transporter pseudogene fragment,
len: 34 aa; highly similar to part of TR:P95156 O08061,
MTCY359.15c, modB, membrane spanning subunit of
ABC-transport system (264 aa), (68.8% identity in 32 aa
overlap)"
/pseudo
/codon_start=1
/transl_table=11
/complement(2418, .2804)
/notes="MNCB1788.09c, ABC-transporter pseudogene fragment,
len:128 aa; highly similar to part of TR:P95157 O05125,
MTCY359.16c, modA, molybdenum-binding subunit of
ABC-transport system (261 aa), (42.3% identity in 97 aa
overlap)"
/pseudo
/codon_start=1
/transl_table=11
/db_xref="PSEUDO:CA15441.1"
/complement(2801, .3058)
/notes="MNCB1788.10c, ABC-transport system pseudogene
fragment, len: 85 aa; highly similar to part of TR:P95157
O05125, MTCY359.16c, modA, molybdenum-binding subunit of
ABC-transport system (261 aa), (61.2% identity in 85 aa
overlap)"
/pseudo
/codon_start=1
/transl_table=11
/notes="possible RBS upstream of MNCB1788.11c"
3300, .3306
/notes="possible RBS upstream of MNCB1788.11c"
3308, .3392
/notes="MNCB1788.11c"
3308, .3392
/notes="MNCB1788.11c"
/notes="MNCB1788.11c, probable oxidoreductase, len: 224 aa;
similar to eg. FABG_ECOLI_P25716 3-oxoacyl-l-acyl-carrier
protein] reductase (244 aa), fasta scores; opt: 157
z-score: 273.8 E(): 4.2e-08, 27.9% identity in 240 aa
overlap. Also highly similar to M. tuberculosis
oxidoreductase MTCY359.17 (80.4% identity in 224 aa
overlap)"
/codon_start=1

Query Match 91.1%; Score 16.4; DB 1; Length 39228;
Best Local Similarity 94.4%; Pred. No. 5,4e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

CY 1 GAGACAAAGAGATTCCG 18
DB 15222 GAGTACAAAGATTCCG 15239

RESULT 4 AC008047 100867 bp DNA linear PLN 28-JUN-2000
LOCUS AC008047
DEFINITION Genomic sequence for Arabidopsis thaliana BAC F2K11 from chromosome
1, complete sequence.
ACCESSION AC008047
VERSION AC008047.3 GI:6623889
KEYWORDS HTG.

```

```

SOURCE
ORGANISM
Arabidopsis thaliana (thale cress)
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 100867)
REFERENCE
AUTHORS
Shinn, P., Brookes, S., Buehler, E., Chao, O., Johnson-Hopson, C.,
Khan, S., Kim, C., Altafi, H., Bei, O., Chin, C., Chiu, J., Choi, E.,
Conn, L., Conway, A., Gonzales, A., Hansen, N., Howing, B., Koo, T.,
Lam, B., Lee, U., Lenz, C., Li, U., Liu, A., Liu, K., Liu, S.,
Mukharasky, N., Nguyen, M., Palm, C., Pham, P., Sakano, H.,
Southwick, A., Thaveri, A., Toriumi, M., Vayberg, M., Yu, G.,
Federpspiel, N.A., Theologis, A. and Ecker, J.R.
Genomic sequence for Arabidopsis thaliana BAC F2K11 from chromosome
1
1 Unpublished
2 (bases 1 to 100867)
REFERENCE
AUTHORS
Ecker, J.R.
Direct Submission
Submitted (15-JUN-1999) Arabidopsis thaliana Genome Center,
Department of Biology, University of Pennsylvania, 38th Street and
Hamilton Walk, Philadelphia, Pennsylvania 19104-6018, USA
3 (bases 1 to 100867)
REFERENCE
AUTHORS
Ecker, J.R.
Direct Submission
Submitted (21-DEC-1999) Arabidopsis thaliana Genome Center,
Department of Biology, University of Pennsylvania, 38th Street and
Hamilton Walk, Philadelphia, Pennsylvania 19104-6018, USA
4 (bases 1 to 100867)
REFERENCE
AUTHORS
Shinn, P., Brookes, S., Buehler, E., Chao, O., Johnson-Hopson, C.,
Khan, S., Kim, C., Altafi, H., Bei, O., Chin, C., Chiu, J., Choi, E.,
Conn, L., Conway, A., Gonzales, A., Hansen, N., Howing, B., Koo, T.,
Lam, B., Lee, U., Lenz, C., Li, U., Liu, A., Liu, K., Liu, S.,
Mukharasky, N., Nguyen, M., Palm, C., Pham, P., Sakano, H.,
Southwick, A., Thaveri, A., Toriumi, M., Vayberg, M., Yu, G., Davis, R.,
Federpspiel, N.A., Theologis, A. and Ecker, J.R.
Direct Submission
Submitted (23-DEC-1999) Arabidopsis thaliana Genome Center,
Department of Biology, University of Pennsylvania, 38th and
Hamilton Walk, Philadelphia, PA 19104-6018, USA
5 (bases 1 to 100867)
REFERENCE
AUTHORS
Cheuk, R., Shinn, P., Brookes, S., Buehler, E., Chao, O.,
Johnson-Hopson, C., Khan, S., Kim, C., Altafi, H., Bei, O.,
Chiu, J., Choi, E., Conn, L., Conway, A., Gonzales, A., Hansen, N.,
Howing, B., Koo, T., Lam, B., Lee, U., Lenz, C., Li, U., Liu, A., Liu, J.,
Liu, S., Mukharasky, N., Nguyen, M., Palm, C., Pham, P., Sakano, H.,
Schwarz, U., Southwick, A., Thaveri, A., Toriumi, M., Vayberg, M.,
Yu, G., Davis, R., Federpspiel, N., Theologis, A. and Ecker, J.R.
Direct Submission
Submitted (28-JUN-2000) Arabidopsis thaliana Genome Center,
Department of Biology, University of Pennsylvania, 38th and
Hamilton Walk, Philadelphia, PA 19104-6018, USA
On Dec 21, 1999 this sequence version replaced gi:5670168.
COMMENT
FEATURES
source
1..100867
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/db_xref="taxon:3702"
/chromosome="1"
/clone="F2K11"
/contig="F2K11"
join(1694, .1814, 2131, .2282, 2395, .2493, 2874, .3083,
3173, .3256, 3400, .3477, 3611, .3668, 4074, .4415, 4496,
4675, .4770, 4859, .5232, 5318, .5458, 5578, .5825, 5514, .6167,
6274, .6437, 6543, .7129, 7218, .7437)
/notes="similar to kinesin-like protein gi|3392771, similar
to ESTs gb|U60306, gb|A1560436.1, gb|T22003, gb|W43855,
and gb|A585767"
/codon_start=1
/evidence=not experimental
/product="F2K11.1"
/protein_id="AA19694.1"
/db_xref="GI:6633835"
/translation="VSSHLSODANWNGVYVRSDVSVSMLSFDGSEBSRESKMDSKKGHOS

```

LYEWNETLPYKLPWEASEDELRACLRDGYLCSLLNQLSPGSMNGSPFPAVYKI
ERFLTAMDEMALPREVSDIEQLKALNKAESDGSYDKNLSLARMSLPDHSRSG
DDRNFTDQSGESGIDMSDAKISDLKSNLSRAPRSLFMDMLDLKSDMSRNG
HYSHAMASLSLVGVIEORISNADLNKSNKGLRLGSDWNGKRGRTVTSSE
KTRIEKESSEKDVRLKREKERSDARPOLKXKPHENOCLELAKAKOKTR
DELEKKNDAELHVVDSRKYELKELCKOSKORKEKCIYONFIDHNSALDEL
TSLSIKHEVNTORXKTFEDNNTYGLKGVADAAKANYVLEENRRLNVAOLKNI
RYCIRPPLPQONSROTTETIGTGLVWANKPOKDRHLFKPKYVVOZAKOTG
LNLFPSSVYLAASSIYDIHILNLLSVEEFLDTRPLRLILDGYNCIPAYCOTS
GTYTMSGSPISITSKEDGWNVALNDLFLTOSRONTYMEGVOMVEIYEOHNGG
LVLTGFTVLMRSIONFLNLTIGLIGNTALPNGLAPDASMHCVSTEDVLEMTG
LMNRVGTATLNERSSRSHCVLSVHNGVDVETDSILGSHLVLDLGSBRVDSERT
GERLKEAOKHINKSIALGDVIFALAKRPHPRYRSKRTVOYLOSGLGQAKTLMDV
NPDGSYATVSTTLKFAEVSQVELGAASKSKEGDVQMEVSNLKADVIKDEEL
ONFOVKGNMNTSLKRGSLNRLVGPPTSRRHSIGASENARAKASGLFGRTSDVN
CEYSKHSKSDSGSQSDSRKQODHQPSFAGAKGIDPDDEVLVGLADSDSD
RLSDSDSCLMGTETDGSISAVELLIPETAKPLIEREBARMTSEKLEKVGK
KTEPDSTRNIPSKIPKOTLPPQOTRPSRLSNSSSSKXLTGNSIHQFDLCPRN
OICPLTILIGSNHSLRFL"
complement (1672..9535,9606..97401)
/note="hypothetical protein"
/codon_start=1
/evidence=not experimental
/product="F2K11.2"
/protein_id="AAFI9720.1"
/db_xref="GI:6633861"
/translation="MWENGCPNVTENTLNLSICRGREVLBALVDRNVEGHPN
AMGSHIKADVYSTALVDRLCKDGNINMKNLTEMKEGIFPNVLTJNCMDISFC
SGRMSADQQLRMIKQINPDIVFSAIINAFKERVABEIEYEMLRWBIPTT
ITNSMIDGFCQKODVDAKMLDSMAKSGSPVVFSTLINGYCAKADVMGEIF
CEMRGRIVANTVYTLIHGFCQVGLDAADQILNEMISGVAPDYITFHQDLGIC
SKELRAFAILEDLOKSEVCPRLIFPLFSIGLIGFGEVCRIYOVKREKQAC
JOH (11234..11473,11567..11656,11735..11815,11913..12033,
12547..12683)
/note="unknown; similar to ESTs gb|AM032410,
gb|AM053721.1, and gb|A1995565.1"
/codon_start=1
/evidence=not experimental
/product="F2K11.3"
/protein_id="AAFI9695.1"
/db_xref="GI:6633836"
/translation="MCSLSYQVFSILQSPSPRPPSLANHEPLSTSSVTFPLKTN
TKOSGIGNLVLRKAYVSSSDSSADSTPPNGTRPKSRDILLEYVNVNPEFEM
FVKRAPGHVVEAKOTVMIGTLPPEFPAVTVAENIACLMASVLTGNTVGRAC
YRLQLQSLQEVALLPEPRGITLALRKLKMSKISSPWKQSRKSTVWKQENQIRSMK
SWSI"
complement (14209..15126)
/note="similar to receptor-like protein kinase gi|4008012;
similar to ESTs gb|T21898 and gb|R90176"
/codon_start=1
/evidence=not experimental
/product="F2K11.4"
/protein_id="AAFI9719.1"
/db_xref="GI:6633860"
/translation="WATMIOSSLSEFCFSLSLSLFSTNLTISESDHIMSTFCQFS
DNFTQSTYETNRETLSSLRBSLIGSTSNATAGISPTVGMFLCRDIDSETSCSD
CVGTATLISRKTQKKAIFIEECMVAYSDSPFSLVDERPYIRYSLSYAPMLDR
FPGTUSDKWDELINATSPSLSTPYFVEDERVSFVDFGSDIDSMACSPDLDRN
CTTCLKLAVOEMLECCNOSQMAOIFPKCLLYEATLALSSPPPPYSPSPSPSP
LSPSPLEKRPQASGSGSSSINVKGN"
complement (16131..16937)
/note="similar to serine/threonine kinase-like protein
emb|CA18465 similar to EST gb|A1994269.1"
/codon_start=1
/evidence=not experimental
/product="F2K11.5"
/protein_id="AAFI9718.1"
/db_xref="GI:6633859"
/translation="WNGINTHFAVALFCFSSSLRMSQASQMLIDTQPNVTRTSS
YLSNRPTVSTIRNNSISGYSNATAGISPTVGMFLCRDIDSETSCSDVNAATLE
TKSCIFRSALVINSPECTIVRSVNSFTLVEDVSTARFSTGNSLSDSPQFSQTILE
KIDALIFRSALSSLPVYFDQOHVTLGSDYLAHMAVQCSPLDPRNCTVCIRLA
VRLSGCCSHAQFARIFYTKLITYEIALQPNVTSIGVTSSESLVYII"

CDS
complement (18736..19500)
/note="similar to protein kinase-like protein
emb|CA18472 similar to ESTs gb|A1994269.1 and gb|h36750"
/codon_start=1
/evidence=not experimental
/product="F2K11.6"
/protein_id="AAFI9717.1"
/db_xref="GI:6633858"
/translation="WRRITDVLFCHFLFSLCMAQSPONLTDNQPPDTYOSNNTV
LSLNRHSSLGSTYFRALIGLSPNTVYGFLICNTSKTSCSNVASHATLEMDKSES
HDSFVSECVRSYSDNSFSLVDSFATFESYSONDSLSPQFYNQTLPGKLDLIL
KAPSSSPSPVPEVDEKHEVTOVESYDLEMAQCSPLDPSSTCTGLVVEKSEF
CSQSRWARLHFPKCLIRYDISALQCNLTSLGVTYTG"
complement (20529..21542)
/note="similar to kinase-like protein emb|CA18472;
similar to ESTs gb|A1994269.1, gb|T21898, gb|h36750, and
gb|R90176"
/codon_start=1
/evidence=not experimental
/product="F2K11.7"
/protein_id="AAFI9716.1"
/db_xref="GI:6633857"
/translation="WRKVTSTFTYSSSLVLPYVLDITLKKSTKSSFHTNSLSP
LPRKIQNKMAIITLALFFRFFSLSHQMSQPHMTFCSDSFQTSYSTNRN
ILITLSSLSLHYLNATIGLSPDTVGMFLCRDINTTSCDQVQAAIEATNCT
LNKRAFIYDECMVRSVNSFSEFSESPVIVRYSLSRAPSNRPNQTLSSKLDQLP
NVPSTLIPYVEQDERVTOLEGSYDLVSMIQSPDLDPSNCTICRAVATVSTCCG
VPSALIFTPKCIIRYRFLVPSAPSPSSLPISPSSTSPPLSLPQLPPLSPQPPP
LST"
complement (22893..23747)
/note="similar to serine/threonine protein kinase
emb|CA18705.1; similar to ESTs gb|R90176 and
gb|A1994269.1"
/codon_start=1
/evidence=not experimental
/product="F2K11.8"
/protein_id="AAFI9715.1"
/db_xref="GI:6633856"
/translation="WATFIRFTAPLFCFFFLFSLFHSQTMSPQHMATFCDRLASNF
TQTSYXANRETLTSLDRSSLGTYNATIGLSPDTVGMFLCRDITLTSQSDVYK
TATLETITNNCTSRKTALIPYECMVRASVSPFTLVEDGPRVAKVATPSTLSF
SPFOOTLSNRKEOILILIASKSSLSASTPYVVDISVNELEGSYDIDVATVQCSPLD
PANGCVPLRVVKNGLSCCTNARAQYILKCFUKYTTTLPTSQSSGSSSINVMKG
EHNCPPL"
complement (26076..27047)
/note="similar to protein kinase-like protein
emb|CA18472; similar to ESTs gb|T21898 and gb|A1994269.1"
/codon_start=1
Query Match 91.1%; Score 16.4; DB 8; Length 100867;
Best Local Similarity 94.4%; Pred. No. 5.2e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
CY 1 GAGAACCAAGATTCCG 18
DB 69213 GAGAACCAAGATTCCG 69196
RESULT 5
LMFLCHR36_27
WFOCOMMENT
Sequence split into 36 fragments LOCUS LMFLCHR36 Accession AL499624
Fragment Name Begin End
LMFLCHR36_00 1 110000
LMFLCHR36_01 100001 210000
LMFLCHR36_02 200001 310000
LMFLCHR36_03 300001 410000
LMFLCHR36_04 400001 510000
LMFLCHR36_05 500001 610000
LMFLCHR36_06 600001 710000
LMFLCHR36_07 700001 810000
LMFLCHR36_08 800001 910000
LMFLCHR36_09 900001 1010000

LMFLCHR36_10 1000001 1110000
 LMFLCHR36_11 1100001 1210000
 LMFLCHR36_12 1200001 1310000
 LMFLCHR36_13 1300001 1410000
 LMFLCHR36_14 1400001 1510000
 LMFLCHR36_15 1500001 1610000
 LMFLCHR36_16 1600001 1710000
 LMFLCHR36_17 1700001 1810000
 LMFLCHR36_18 1800001 1910000
 LMFLCHR36_19 1900001 2010000
 LMFLCHR36_20 2000001 2110000
 LMFLCHR36_21 2100001 2210000
 LMFLCHR36_22 2200001 2310000
 LMFLCHR36_23 2300001 2410000
 LMFLCHR36_24 2400001 2510000
 LMFLCHR36_25 2500001 2610000
 LMFLCHR36_26 2600001 2710000
 LMFLCHR36_27 2700001 2810000
 LMFLCHR36_28 2800001 2910000
 LMFLCHR36_29 2900001 3010000
 LMFLCHR36_30 3000001 3110000
 LMFLCHR36_31 3100001 3210000
 LMFLCHR36_32 3200001 3310000
 LMFLCHR36_33 3300001 3410000
 LMFLCHR36_34 3400001 3510000
 LMFLCHR36_35 3500001 3529952

Continuation (28 of 36) of LMFLCHR36 from base 2700001 (AL499624 Leishmania major chromo

Query Match 91.1% Score 16.4; DB 2; Length 110000;
 Best Local Similarity 94.4%; Pred. No. 5.2e+02;
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

1 GAGAACACAGATTCCG 18
 |||||
 Db 91900 GAGACCAACAGATTCCG 91917

RESULT 6
 AC143118/c
 LOCUS AC143118 124215 bp DNA linear HTG 09-APR-2003
 DEFINITION Macaca mulatta clone CH250-269J3. *** SEQUENCING IN PROGRESS ***.
 AC143118.1 GI:29567757
 HTG HTGS_PHASE2, HTGS_PGI.
 KEYWORDS Macaca mulatta (rhesus monkey)
 SOURCE Macaca mulatta
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
 Cercopithecinae; Macaca.

REFERENCE
 AUTHORS Csurös, M., and Milosavljević, A.
 TITLE Pooled genomic indexing (PGI): mathematical analysis and experiment
 design
 JOURNAL (in) Guigo, R. and Gusfield, D. (Eds.);
 ALGORITHMS IN BIOINFORMATICS, SECOND INTERNATIONAL WORKSHOP, WABI
 2002, ROME, ITALY, SEPTEMBER 17-21, 2002, PROCEEDINGS: 10-28;
 Springer (2002)

REFERENCE
 AUTHORS 2 (bases 1 to 124215)
 Milosavljević, A., Sodeyren, E., Csurös, M., Li, B., Jackson, A.R.,
 Adams, C., Adio-Oduola, B., Ali-oshman, F.R., Allen, C., Alsdorfs, S.L.,
 Amarantunga, H.C., Are, J.R., Ayale, M., Banks, T., Barbata, J.,
 Beneton, J., Blum, K., Blankenburg, K., Bonnin, D., Bouck, J.,
 Bowie, S., Briviera, M., Brown, E., Bryant, M., Bryant, N.P., Buhay, C.,
 Burch, P., Burkett, C., Burrell, K.L., Byrd, N.C., Carron, T.F.,
 Carter, M., Cavazos, S.R., Chacko, J., Chavez, D., Cleveland, C.D.,
 Chen, Z., Chiu, D., Chowdhury, I., Christopoulos, C., Cleaveland, C.D.,
 Cox, C., Coyle, M.D., Dathorne, S.R., David, R., Davila, M.L., Davis, C.,
 Davy-Carroll, L., Dederich, D.A., Delaney, K.R., Delgado, O.,
 Dem, A.U., Ding, Y., Din, H.H., Douthwaite, K.J., Draper, H.,
 Dugan-Rocha, S., Durbin, K.J., Egan, A., Earmhart, C., Edwards, C.C.,
 Elhai, C., Emerling, S., Escotto, M., Falls, T., Ferraguto, D.,
 Flagg, N., Ford, J., Foster, P., Frantz, P., Gabisi, A., Geo, J.,
 Garcia, A., Garner, T., Garza, N., Gill, R., Gotrell, V.H., Guevara, W.,

TITLE
 JOURNAL
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL

REFERENCE
 AUTHORS
 TITLE
 JOURNAL

COMMENT

Center: Baylor College of Medicine
 Center code: BCM
 Web site: <http://www.hgsc.bcm.tmc.edu/>
 Contact: hgsc-help@bcm.tmc.edu
 ----- Project Information
 Center project name: LNPC
 Center clone name: CH250-269J3
 ----- Summary Statistics
 Chemistry: Dye-primer Bodipy: inf of reads
 Chemistry: Dye-terminator Big Dye: inf of reads
 Consensus quality: 6461 bases at least Q40
 Consensus quality: 7875 bases at least Q30
 Consensus quality: 9630 bases at least Q20

 * NOTE: Estimated insert size may differ from sequence length
 * (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html)
 * NOTE: The contigs are based on the application
 * of the PGI method using the Human genome (NCBI build 31)
 * as the comparative genome.
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 1 contigs. Gaps between the contigs
 * are represented as runs of N. The order of the pieces
 * is believed to be correct as given, however the sizes
 * of the gaps between them are based on estimates that have
 * provided by the submitter.
 * This sequence will be replaced
 * by the finished sequence as soon as it is available and
 * the accession number will be preserved.
 * 1 124215: contig of 124215 bp in length.
 Location/Qualifiers
 1. 124215

FEATURES

source


```

LTR                                     /transposon="LTR retrotransposon, copia, WIS_107G22-2"
                                     complement(43213..44660)
                                     /note="3' LTR of WIS_107G22-2"
LTR                                     complement(48849..51570)
                                     /note="5' LTR of WIS_107G22-2"
LTR                                     complement(55446..57167)
                                     /note="5' LTR of WIS_107G22-1"
LTR                                     58235..59805
                                     /note="3' LTR of Sabrina_107G22-4"
                                     order(59960..59964,63215..63219)
                                     /note="target site duplication of Wilma_107G22-1"
LTR                                     /rpt_type=direct
                                     join(59965..61261,62989..63214)
                                     /note="Solo-LTR; LTR retrotransposon, gypsy,
                                     Wilma_107G22-1"
                                     order(61257..61261,62989..62993)
                                     /note="target site duplication of WIS_107G22-3"
repeat_region                       /rpt_type=direct
                                     complement(61262..62988)
                                     /note="Solo-LTR; LTR retrotransposon, copia, WIS_107G22-3"
                                     complement(64272..64278)
                                     /note="putative"
                                     complement(64353..65519)
                                     /gene="Glu-A3-1"
                                     complement(64353..65519)
                                     /gene="Glu-A3-1"
                                     /product="TbGlu-A3-1"
                                     complement(64353..65519)
                                     /gene="Glu-A3-1"
                                     /note="low molecular weight glutenin, TdGlu-A3-1"
                                     /codon_start=1
                                     /product="TbGlu-A3-1"
                                     /protein_id="TAOI17157.1"
                                     /db_xref="GI:30349131"
                                     /translation="MKTFILVFAALAAAGAVQISGQQQPPFLQQQQPPFSQQQQP
                                     PFSSQQQSPFQQQQPPFSQQQQPPFSQQQQPPFSQQQQPPFSQQQQPPFSQQQQPPFSQQQQPPFS
                                     QQQQPPFSQQQQPPFSQQQQPPFSQQQQPPFSQQQQPPFSQQQQPPFSQQQQPPFSQQQQPPFS
                                     QQQQPPFSQQQQPPFSQQQQPPFSQQQQPPFSQQQQPPFSQQQQPPFSQQQQPPFSQQQQPPFS
                                     NPKVFIQQQCIPIVAVQRCARSMLOQSSCHVQRCQQCQICQIIPQSHSIRALI
                                     YSILIQQQQQQQQQQQQQQSSITQIQPPQGLGCVSPQQLAQQLGQQPQQQQ
                                     LAHGTFIQPIQIAQLFWMTSIALKTLPRMGSVNVPIETTSVPLGVGIGVGY"
repeat_region                       64633..64661
                                     /note="microsatellite"
                                     /rpt_type=direct
                                     complement(65576..65581)
                                     /rpt_unit="ctg"
                                     /rpt_type=direct
                                     complement(65593..65598)
                                     /note="putative TATA_signal-1"
                                     67466..69851
                                     /note="region of RFLP marker SFR159"
                                     order(70504..70506,71431..71433)
                                     /note="target site duplication of Mandrake_107G22-1"
                                     /rpt_type=direct
                                     70507..71430
                                     /transposon="transposon, CACTA, Mandrake_107G22-1; SNAC
                                     (Small non-autonomous CACTA) transposon"
                                     77451..78390
                                     /gene="LRR-1"
                                     /pseudo
                                     join(77451..77807,77809..78088,78275..78390)
                                     /gene="LRR-1"
                                     /note="TLRR-1; NBS-LRR resistance gene analog; TdLRR-1"
                                     /pseudo
                                     /codon_start=1
                                     order(79798..79800,80073..80075)
                                     /note="target site duplication of Orpheus_107G22-2"
                                     /rpt_type=direct
                                     79801..80072

```

```

Query Match          91.1%; Score 16.4; DB 8; Length 142018;
Best Local Similarity 94.4%; Pred. No. 5.2e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 GAGAACACACAGATTTCGC 18
DB      68155 GAGCACACAGATTTCGC 68138

RESULT 8
CR352284/c
LOCUS      CR352284
DEFINITION Danio rerio clone DKer-89F23, *** SEQUENCING IN PROGRESS ***, 14
unordered pieces.
ACCESSION CR352284
VERSION   CR352284.4 GI:45433392
KEYWORDS  HTG; HTGS PHASE1.
SOURCE    Danio rerio (zebrafish)
ORGANISM  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Danio.
REFERENCE 1 (bases 1 to 169261)
AUTHORS   McIay, K.
TITLES    Direct Submission
JOURNAL   Submitted (10-MAR-2004) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
clonerequests@sanger.ac.uk
On Mar 13, 2004 this sequence version replaced gi:45381851.

COMMENT
----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: zfish-help@sanger.ac.uk
----- Project Information
Center project name: zK89F23
----- Summary Statistics
Assembly program: XGAP4; version 4.5
Chemistry: Dye-termination; 100% of reads
Consensus quality: 164159 bases at least Q40
Consensus quality: 165359 bases at least Q30
Consensus quality: 166499 bases at least Q20
Insert size: 167961; sum-of-contigs
Insert size: 186476; 1.3% error; agarose-fp
Quality coverage: 7.82x in Q20 bases; sum-of-contigs Quality
coverage: 7.35x in Q20 bases; agarose-fp
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 14 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1
11905: contig of 11905 bp in length
11906: gap of 100 bp
12005: gap of 100 bp
12006: contig of 2615 bp in length
14621: gap of 100 bp
14720: gap of 100 bp
14721: gap of 14241 bp in length
14722: contig of 100 bp
14723: gap of 100 bp
14724: gap of 100 bp
14725: gap of 100 bp
14726: gap of 100 bp
14727: gap of 100 bp
14728: gap of 100 bp
14729: gap of 100 bp
14730: gap of 100 bp
14731: gap of 100 bp
14732: gap of 100 bp
14733: gap of 100 bp
14734: gap of 100 bp
14735: gap of 100 bp
14736: gap of 100 bp
14737: gap of 100 bp
14738: gap of 100 bp
14739: gap of 100 bp
14740: gap of 100 bp
14741: gap of 100 bp
14742: gap of 100 bp
14743: gap of 100 bp
14744: gap of 100 bp
14745: gap of 100 bp
14746: gap of 100 bp
14747: gap of 100 bp
14748: gap of 100 bp
14749: gap of 100 bp
14750: gap of 100 bp
14751: gap of 100 bp
14752: gap of 100 bp
14753: gap of 100 bp
14754: gap of 100 bp
14755: gap of 100 bp
14756: gap of 100 bp
14757: gap of 100 bp
14758: gap of 100 bp
14759: gap of 100 bp
14760: gap of 100 bp
14761: gap of 100 bp
14762: gap of 100 bp
14763: gap of 100 bp
14764: gap of 100 bp
14765: gap of 100 bp
14766: gap of 100 bp
14767: gap of 100 bp
14768: gap of 100 bp
14769: gap of 100 bp
14770: gap of 100 bp
14771: gap of 100 bp
14772: gap of 100 bp
14773: gap of 100 bp
14774: gap of 100 bp
14775: gap of 100 bp
14776: gap of 100 bp
14777: gap of 100 bp
14778: gap of 100 bp
14779: gap of 100 bp
14780: gap of 100 bp
14781: gap of 100 bp
14782: gap of 100 bp
14783: gap of 100 bp
14784: gap of 100 bp
14785: gap of 100 bp
14786: gap of 100 bp
14787: gap of 100 bp
14788: gap of 100 bp
14789: gap of 100 bp
14790: gap of 100 bp
14791: gap of 100 bp
14792: gap of 100 bp
14793: gap of 100 bp
14794: gap of 100 bp
14795: gap of 100 bp
14796: gap of 100 bp
14797: gap of 100 bp
14798: gap of 100 bp
14799: gap of 100 bp
14800: gap of 100 bp
14801: gap of 100 bp
14802: gap of 100 bp
14803: gap of 100 bp
14804: gap of 100 bp
14805: gap of 100 bp
14806: gap of 100 bp
14807: gap of 100 bp
14808: gap of 100 bp
14809: gap of 100 bp
14810: gap of 100 bp
14811: gap of 100 bp
14812: gap of 100 bp
14813: gap of 100 bp
14814: gap of 100 bp
14815: gap of 100 bp
14816: gap of 100 bp
14817: gap of 100 bp
14818: gap of 100 bp
14819: gap of 100 bp
14820: gap of 100 bp
14821: gap of 100 bp
14822: gap of 100 bp
14823: gap of 100 bp
14824: gap of 100 bp
14825: gap of 100 bp
14826: gap of 100 bp
14827: gap of 100 bp
14828: gap of 100 bp
14829: gap of 100 bp
14830: gap of 100 bp
14831: gap of 100 bp
14832: gap of 100 bp
14833: gap of 100 bp
14834: gap of 100 bp
14835: gap of 100 bp
14836: gap of 100 bp
14837: gap of 100 bp
14838: gap of 100 bp
14839: gap of 100 bp
14840: gap of 100 bp
14841: gap of 100 bp
14842: gap of 100 bp
14843: gap of 100 bp
14844: gap of 100 bp
14845: gap of 100 bp
14846: gap of 100 bp
14847: gap of 100 bp
14848: gap of 100 bp
14849: gap of 100 bp
14850: gap of 100 bp
14851: gap of 100 bp
14852: gap of 100 bp
14853: gap of 100 bp
14854: gap of 100 bp
14855: gap of 100 bp
14856: gap of 100 bp
14857: gap of 100 bp
14858: gap of 100 bp
14859: gap of 100 bp
14860: gap of 100 bp
14861: gap of 100 bp
14862: gap of 100 bp
14863: gap of 100 bp
14864: gap of 100 bp
14865: gap of 100 bp
14866: gap of 100 bp
14867: gap of 100 bp
14868: gap of 100 bp
14869: gap of 100 bp
14870: gap of 100 bp
14871: gap of 100 bp
14872: gap of 100 bp
14873: gap of 100 bp
14874: gap of 100 bp
14875: gap of 100 bp
14876: gap of 100 bp
14877: gap of 100 bp
14878: gap of 100 bp
14879: gap of 100 bp
14880: gap of 100 bp
14881: gap of 100 bp
14882: gap of 100 bp
14883: gap of 100 bp
14884: gap of 100 bp
14885: gap of 100 bp
14886: gap of 100 bp
14887: gap of 100 bp
14888: gap of 100 bp
14889: gap of 100 bp
14890: gap of 100 bp
14891: gap of 100 bp
14892: gap of 100 bp
14893: gap of 100 bp
14894: gap of 100 bp
14895: gap of 100 bp
14896: gap of 100 bp
14897: gap of 100 bp
14898: gap of 100 bp
14899: gap of 100 bp
14900: gap of 100 bp
14901: gap of 100 bp
14902: gap of 100 bp
14903: gap of 100 bp
14904: gap of 100 bp
14905: gap of 100 bp
14906: gap of 100 bp
14907: gap of 100 bp
14908: gap of 100 bp
14909: gap of 100 bp
14910: gap of 100 bp
14911: gap of 100 bp
14912: gap of 100 bp
14913: gap of 100 bp
14914: gap of 100 bp
14915: gap of 100 bp
14916: gap of 100 bp
14917: gap of 100 bp
14918: gap of 100 bp
14919: gap of 100 bp
14920: gap of 100 bp
14921: gap of 100 bp
14922: gap of 100 bp
14923: gap of 100 bp
14924: gap of 100 bp
14925: gap of 100 bp
14926: gap of 100 bp
14927: gap of 100 bp
14928: gap of 100 bp
14929: gap of 100 bp
14930: gap of 100 bp
14931: gap of 100 bp
14932: gap of 100 bp
14933: gap of 100 bp
14934: gap of 100 bp
14935: gap of 100 bp
14936: gap of 100 bp
14937: gap of 100 bp
14938: gap of 100 bp
14939: gap of 100 bp
14940: gap of 100 bp
14941: gap of 100 bp
14942: gap of 100 bp
14943: gap of 100 bp
14944: gap of 100 bp
14945: gap of 100 bp
14946: gap of 100 bp
14947: gap of 100 bp
14948: gap of 100 bp
14949: gap of 100 bp
14950: gap of 100 bp
14951: gap of 100 bp
14952: gap of 100 bp
14953: gap of 100 bp
14954: gap of 100 bp
14955: gap of 100 bp
14956: gap of 100 bp
14957: gap of 100 bp
14958: gap of 100 bp
14959: gap of 100 bp
14960: gap of 100 bp
14961: gap of 100 bp
14962: gap of 100 bp
14963: gap of 100 bp
14964: gap of 100 bp
14965: gap of 100 bp
14966: gap of 100 bp
14967: gap of 100 bp
14968: gap of 100 bp
14969: gap of 100 bp
14970: gap of 100 bp
14971: gap of 100 bp
14972: gap of 100 bp
14973: gap of 100 bp
14974: gap of 100 bp
14975: gap of 100 bp
14976: gap of 100 bp
14977: gap of 100 bp
14978: gap of 100 bp
14979: gap of 100 bp
14980: gap of 100 bp
14981: gap of 100 bp
14982: gap of 100 bp
14983: gap of 100 bp
14984: gap of 100 bp
14985: gap of 100 bp
14986: gap of 100 bp
14987: gap of 100 bp
14988: gap of 100 bp
14989: gap of 100 bp
14990: gap of 100 bp
14991: gap of 100 bp
14992: gap of 100 bp
14993: gap of 100 bp
14994: gap of 100 bp
14995: gap of 100 bp
14996: gap of 100 bp
14997: gap of 100 bp
14998: gap of 100 bp
14999: gap of 100 bp
15000: gap of 100 bp

```

FEATURES	
*	112706 contig of 2743 bp in length
*	114748 contig of 100 bp
*	114749 contig of 2795 bp in length
*	114848 contig of 100 bp
*	114849 contig of 2795 bp in length
*	136643 contig of 100 bp
*	136743 contig of 100 bp
*	151727 contig of 1984 bp in length
*	151827 contig of 100 bp
*	151828 contig of 2235 bp in length
*	151829 contig of 2235 bp in length
*	154062 contig of 100 bp
*	154162 contig of 100 bp
*	154163 contig of 3308 bp in length
*	157470 contig of 100 bp
*	157471 contig of 100 bp
*	157570 contig of 100 bp
*	157571 contig of 1691 bp in length
	Location/Qualifiers

ACCESSION CR382377
VERSION CR382377.3 GI:46241759
KEYWORDS HTG; HTGS_Phasel.
SOURCE Danio rerio (zebrafish)
ORGANISM Danio rerio
Eukaryotes; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes; Cyprinidae; Danio.
1 (bases 1 to 214536)
McLay, K.
REFERENCE
AUTHORS
TITLE Direct Submission
JOURNAL Submitted (01-APR-2004) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
zish-help@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
COMMENT On Apr 6, 2004 this sequence version replaced gi:46241593.

COMMENT

1. 169261

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc feature

misc feature

misc feature

misc_feature

 \mathbb{Z}

Local Similarity Match

1 GAG1

5659 GTGA

9
T
9

MITION Danio

CR382377 214536 bp DNA linear HTG 03-APR-2004
 Danio rerio clone DKEY-238113, *** SEQUENCING IN PROGRESS ***, 28
 unordered pieces.

CR382377
CR382377.3 GI:462411759
HTG; HTGS PHASL1.
Danio rerio (*zebrafish*)
Danio rerio
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neoceratognathi; Teleostei; Ostariophysi; Cypriniformes; Cyprinidae; Danio.
1 (bases 1 to 214536)
McLay,K.

Direct Submission
Submitted (01-MAR-2004) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
zfish-hlps@sanger.ac.uk
On Apr 6, 2004 this sequence version replaced gi:46241599.

----- Genome Center -----
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: zfish-hlps@sanger.ac.uk
Project information
----- Project Information -----
Center project name: ZK38113
Summary Statistics
Assembly program: XCAP4, version 4.5
Chemistry: Dye-terminator, 100% of reads
Consensus quality: 205798 bases at least Q40
Consensus quality: 207592 bases at least Q30
Consensus quality: 208725 bases at least Q20
Insert size: 21836; sum-of-contigs
Insert size: 170027; 7.1% error; agarose-fp
Quality coverage: 5.74x in Q20 bases; sum-of-contigs Quality coverage: 8.57x in Q20 bases; agarose-fp

* NOTE: This is a 'working draft' sequence. It currently consists of 28 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

	1	24042	contig of 24042 bp in length
*	24043	24142	gap of 100 bp
*	24143	44511	contig of 22469 bp in length
*	44512	44511	gap of 100 bp
*	4512	50349	contig of 4538 bp in length
*	50250	50349	gap of 100 bp
*	50350	55662	contig of 6613 bp in length
*	55663	57062	gap of 100 bp
*	57063	70485	contig of 13423 bp in length
*	70486	89747	gap of 100 bp
*	70586	89747	contig of 19162 bp in length
*	89748	89847	gap of 100 bp
*	89848	93674	contig of 3827 bp in length
*	93675	93774	gap of 100 bp
*	93775	96584	contig of 2810 bp in length
*	96585	96684	gap of 100 bp
*	96685	107986	contig of 10712 bp in length
*	10797	107996	gap of 100 bp
*	107497	145668	contig of 38172 bp in length
*	145669	145668	gap of 100 bp
*	145769	158179	contig of 12411 bp in length
*	158180	158279	gap of 100 bp
*	158280	162104	contig of 3825 bp in length
*	162105	162204	gap of 100 bp
*	162205	169242	contig of 6938 bp in length
*	169143	169242	gap of 100 bp
*	169243	171568	contig of 2326 bp in length
*	171569	171668	gap of 100 bp
*	171669	175613	contig of 3945 bp in length
*	175614	175713	gap of 100 bp
*	175714	178967	contig of 3254 bp in length
*	178968	179067	gap of 100 bp

```

* 179068 181118: contig of 2051 bp in length
* 181119 181218: gap of 100 bp
* 181219 18370: contig of 2252 bp in length
* 18371 18370: gap of 100 bp
* 18371 185728: contig of 2158 bp in length
* 185729 185828: gap of 100 bp
* 185829 188116: contig of 2288 bp in length
* 188117 188216: gap of 100 bp
* 188217 190461: contig of 2245 bp in length
* 190462 190561: gap of 100 bp
* 190562 193442: contig of 2881 bp in length
* 193443 193542: gap of 100 bp
* 193543 19057: contig of 5515 bp in length
* 19058 19157: gap of 100 bp
* 19158 202015: contig of 2858 bp in length
* 202016 202115: gap of 100 bp
* 202116 204305: contig of 2190 bp in length
* 204306 204405: gap of 100 bp
* 204406 206588: contig of 2183 bp in length
* 206589 206688: gap of 100 bp
* 206689 209849: contig of 3161 bp in length
* 209850 209949: gap of 100 bp
* 209950 214536: contig of 4587 bp in length.
FEATURES
    source
        1..214536
            /organism="Danio rerio"
            /mol_type="genomic DNA"
            /db_xref="taxon:7955"
            /clone_lib="DKEY-238113"
            /clone_libs="DanioKey"
        1..24042
            /note="assembly fragment:00928"
            /fragment_chain:1
            24143..45611
            /note="assembly fragment:01235"
            /fragment_chain:1
            45712..50249
            /note="assembly fragment:00001"
            /fragment_chain:1
            50350..56962
            /note="assembly fragment:00166"
            /fragment_chain:1
            57083..70485
            /note="assembly fragment:00552"
            /fragment_chain:1
            70586..89747
            /note="assembly fragment:00739"
            /fragment_chain:1
            89848..93674
            /note="assembly fragment:02435"
            /fragment_chain:1
            93775..96584
            /note="assembly fragment:02387"
            /fragment_chain:1
            96685..107396
            /note="assembly fragment:00267"
            /fragment_chain:1
            107497..145668
            /note="assembly fragment:01551"
            /fragment_chain:1
            145769..158179
            /note="assembly fragment:00402"
            /fragment_chain:1
            158280..162104
            /note="assembly fragment:02467"
            /fragment_chain:1
            162205..169142
            /note="assembly fragment:00073"
            /fragment_chain:1
            169243..171568
            /note="assembly fragment:02257"
            /fragment_chain:2
            171669..175613

```

```

/note="assembly fragment:02401"
/fragment_chain:2
175714..178967
/note="assembly fragment:02375"
/fragment_chain:2
179068..181118
/note="assembly fragment:02343"
/fragment_chain:2
181219..183470
/note="assembly fragment:02292"
/fragment_chain:3
183571..185728
/note="assembly fragment:02325"
/fragment_chain:3
185829..188116
/note="assembly fragment:02307"
/fragment_chain:4
188217..190461
/note="assembly fragment:02313"
/fragment_chain:4
190562..193442
/note="assembly fragment:02357"
/fragment_chain:5
193543..199057
/note="assembly fragment:02417"
/fragment_chain:5
199158..202015
/note="assembly fragment:02277"
202116..204305
/note="assembly fragment:02331"
204406..206588
/note="assembly fragment:02350"
206689..209849
/note="assembly fragment:02365"
209950..214536
/note="assembly fragment:02501.0"

```

ORIGIN

```

Query Match      91.1%; Score 16.4; DB 2; Length 214536;
Best Local Similarity 94.4%; Pred. No. 5.1e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

QY 1 GAGAACACACAGATTCCG 18

DB 209027 GTGAACACACAGATTCCG 209044

RESULT 10

AC111804 AC111804 239570 bp DNA linear HTG 15-NOV-2002
 LOCUS Rattus norvegicus clone CH230-27013, WORKING DRAFT SEQUENCE.

DEFINITION AC111804
 AC111804
 AC111804.4 GI:25006998
 HTG; HTGS PHASE2; HTGS DRAFT; HTGS_FULLTOP.
 KEYWORDS Rattus norvegicus (Norway rat)
 SOURCE Rattus norvegicus

ORGANISM Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.

REFERENCE 1 (bases 1 to 239570)
 AUTHORS Muzny,D,Marie., Metzker,M, Lee., Abramson,S., Adams,C., Alder,J.,
 Allen,C., Allen,H., Altschrocks,S., Amin,A., Anguiano,D.,
 Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,
 Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
 Biswal,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
 Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,
 Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A.,
 Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
 Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
 Davila,M.B., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
 Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,
 Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,
 Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,

Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Frazer, C.M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M., Gregor, G., Greer, K., Gill, R., Grady, M., Guerra, M., Guevara, W., Gunaratne, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K., Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S.L., Hodgson, A., Hughes, M., Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A., Karpatis, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kowis, C., Kraft, C.L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lorensu, H., Lounsbury, H., Lozano, R.J., Lu, X., Ma, J., Maheshwari, M., Mahindrasekhar, M., Mahmoud, M., Malloy, K., Mangum, A., Mangum, B., Mapus, P., Martin, K., Martin, R., Martinez, E., Mawhinney, S., McLeod, M.P., McNeill, T.Z., Meenen, E., Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwankwelu, O., Okwuonu, G., Olarnpunsagoon, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C., Plopper, F., Poindexter, A., Popovic, D., Primus, E., Pu, L., Puzo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M.A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S., Sanders, M., Savary, G., Scherer, S., Scott, G., Shatsman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C.D., Smajic, D., Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J., Steimle, M., Strong, R., Sutton, A., Svatek, A., Taber, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K., Valas, R., Vera, V., Villalana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wleczky, R., Woodin, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Niederhausen, A., Weiser, R., Smith, D.R., Holt, R.A., Smith, H.O., Weinstock, G., and Gibbs, R.A.

Direct Submission
Unpublished
2 (bases 1 to 239570)
Worley, K.C.
Direct Submission
Submitted (19-FEB-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 239570)
Rat Genome Sequencing Consortium.
Direct Submission
Submitted (15-NOV-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
On Nov 15, 2002 this sequence version replaced gi:23195895.
The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL

COMMENT

Center: Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GOSE
Center clone name: CH230-27013
----- Summary Statistics
Assembly program: Phrap; version 0.990329

Consensus quality: 227212 bases at least Q40
Consensus quality: 228362 bases at least Q30
Consensus quality: 229227 bases at least Q20
Estimated insert size: 233057; sum-of-contigs estimation
Quality coverage: 9x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 1 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
* 1 239570: contig of 239570 bp in length.
* Location/Qualifiers
1..239570
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-27013"
1..1150
/note="wgs end extension
clone_end:T7"
complement(7684..8310)
/note="clone boundary
clone_end:T7
site:
end_sequence:BH338730"
complement(238848..239097)
/note="clone boundary
clone_end:Sp6
site:
end_sequence:BH338732"
ORIGIN
Query Match 91.1%; Score 16.4; DB 2; Length 239570;
Best Local Similarity 94.4%; Pred. No. 5.1e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 GAGAACACACAGATTGCG 18
DB 111715 GAGAACACACAGATTGCG 111732

RESULT 11
LOCUS
AC111225/c
DEFINITION
Rattus norvegicus clone CH230-109120, *** SEQUENCING IN PROGRESS
***, 3 unordered pieces.
AC111225
AC111225.5 GI:30578498
HTG; HTGS PHASE1; HTGS DRAFT; HTGS_ENRICHED.
KEYWORDS
Rattus norvegicus (Norway rat)
SOURCE
Rattus norvegicus
ORGANISM
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 256123)
Muzny, D., Maric, M., Metzker, M., Lee, S., Adams, C., Alder, J., Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguiano, D., Anyalebechi, V., Ayvaci, A., Ayodeji, M., Baca, E., Baden, H., Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benham, P., Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M., Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E., Cardenas, V., Carter, K., Cavazos, I., Cesar, H., Center, A., Chacko, J., Chavez, D., Chen, G., Chen, Y., Chen, Z., Chu, J., Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L., Davila, M.L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D., Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K.,

Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K., Egan, A., Escotto, M., Eugene, C., Evans, C.A., Falls, I., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Fraser, C.M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M., Gebregeorgis, E., Geier, K., Gill, R., Grady, M., Guerra, T., Guevara, W., Gunaratne, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K., Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S.L., Hodgson, A., Hogues, M., Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A., Karpathy, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kowis, C., Kraft, C.L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, J., Liu, Y., London, P., Longacre, S., Lopez, J., Lorusshewa, L., Loulseghe, H., Lozado, R.J., Lu, X., Ma, J., Maheshwari, M., Mahindaratne, M., Mahmoud, M., Malloy, K., Mangum, A., Mangum, B., Mapa, P., Martin, K., Martin, R., Martinez, E., Mahoney, S., McGeod, M.P., McNeill, T.Z., Meenen, E., Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Munidas, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwackelmele, O., Okwunu, G., Olarunpunsagcon, A., Pal, S., Parks, K., Pasternak, F., Paul, H., Perez, A., Perez, L., Pfannkuch, C., Plopper, F., Poindexter, A., Popovic, D., Primus, E., Pu, L.-L., Puazo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M.A., Reich, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S.J., Sanders, W., Savary, G., Scherer, S., Scott, G., Shatsman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C.D., Smajs, D., Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J., Steidle, M., Strong, R., Sutton, A., Svatek, A., Taber, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K., Valas, R., Vera, V., Villanar, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wlezyk, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Niederhausen, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O., Weinstock, G. and Gibbs, R.A.

Direct Submission
Unpublished
2 (bases 1 to 256123)
Worley, K.C.

Submitted (19-FEB-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

3 (bases 1 to 256123)
Rat Genome Sequencing Consortium.

Direct Submission
Submitted (13-MAY-2003) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

On May 13, 2003 this sequence version replaced gi:24942391. The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GLVI
Center clone name: CH230-109120

----- Summary Statistics
Assembly program: Atlas 3.0;
Consensus quality: 237968 bases at least Q40
Consensus quality: 240525 bases at least Q30
Consensus quality: 242121 bases at least Q20
Estimated insert size: 256749; sum-of-contigs estimation
Quality coverage: 6x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/GCSC/genbank_draft_data.html)
* NOTE: This sequence may represent more than one clone.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 3 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence.
* as soon as it is available and the accession number will
* be preserved.

1 251516: contig of 251516 bp in length
251517 251616: gap of unknown length
251617 253761: contig of 2145 bp in length
253762 253861: gap of unknown length
253862 256123: contig of 2262 bp in length.

Location/Qualifiers
1. 256123
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-109120"

1. 2094
/note="wgs_end_extension
clone_end:77"
3849. 4622
/note="clone_boundary
clone_end:T7
site:ECORI
end sequence:BH300712"
175598. 176593
/note="wgs_contig"
246978. 247860
/note="clone_boundary
clone_end:Sp6
site:ECORI
end sequence:BH300713"
247056. 247857
/note="clone_boundary
clone_end:Sp6
site:ECORI
end sequence:BH300713"
248853. 251516
/note="wgs_end_extension
clone_end:Sp6"

FEATURES
source
misc_feature
misc_feature
misc_feature
misc_feature
misc_feature
misc_feature

ORIGIN
Query Match 91.1%; Score 16.4; DB 2; Length 256123;
Best Local Similarity 94.4%; Pred. No. 5.1e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GAGAACACACAGATTCCG 18
|||||
DB 148114 GAGAACACACAGATTCC 148037
|||||

RESULT 12
MLEPRTN8
LOCUS MLEPRTN8 342300 bp DNA linear BCT 22-FEB-2001
DEFINITION Mycobacterium leprae strain TN complete genome; segment 8/10.
ACCESSION AL583924 AL50380
VERSION AL583924.1 GI:13093618
KEYWORDS Mycobacterium leprae
SOURCE Mycobacterium leprae
ORGANISM Mycobacterium leprae

Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.

1 (bases 1 to 342300)
 Cole, S.T., Eiglmeyer, K., Parkhill, J.J., James, K.D., Thomson, N.R., Wheeler, P.R., Honore, N., Ganier, T., Churcher, C., Harris, D., Mungall, K., Basham, D., Brown, D., Chillingworth, T., Connor, R., Davies, R.M., Devlin, K., Duthoy, S., Feltwell, T., Fraser, A., Hamlin, N., Holroyd, S., Hornsby, T., Jagels, K., Lacroix, C., Maclean, J., Moule, S., Murphy, L., Oliver, Quail, M.A., Rajandream, M.-A., Rutherford, K.M., Rutter, S., Seeger, K., Simon, S., Simmonds, M., Skelton, J., Squares, R., Stevens, K., Taylor, K., Whitehead, S., Woodward, J.R. and Barrell, B.G.
 Massive gene decay in the leprosy bacillus
 Nature 409 (6823), 1007-1011 (2001)
 21128732
 11234002
 2 (bases 1 to 342300)
 Parkhill, J.J.
 Direct Submission
 Submitted (20-Feb-2001) Submitted on behalf of the Mycobacterium leprae sequencing teams, The Sanger Centre, Wellcome Trust Genome Campus, Hinxton, Cambridge, CB10 1SA, UK Unite de Genetique Moleculaire Bacterienne, Institut Pasteur 28 rue du Docteur Roux, 75724, Paris Cedex, France. E-mail: parkhill@sanger.ac.uk

Notes:
 Details of M. leprae sequencing at the Sanger Centre are available from http://www.sanger.ac.uk/projects/M_leprae/. A relational database containing the M. leprae sequences is available from <http://genolist.pasteur.fr/Leproma/>.

Location/Qualifiers
 1. 342300
 /organism="Mycobacterium leprae"
 /mol_type="genomic DNA"
 /strain="TN"
 /db_xref="taxon:1769"
 275..430
 /gene="ML1963"
 /pseudo
 /pseudo
 275..430
 /gene="ML1963"
 /note="Possible pseudogene of M. tuberculosis orthologue Rv3463 (Best blastx score 121)"
 /pseudo
 /transl_table=11
 /codon_start=1
 /product="probable neuraminidase (pseudogene)"
 483..1484
 /gene="rmlB"
 /note="synonym: ML1964"
 483..1484
 /gene="rmlB"
 /note="Similar to M. tuberculosis dTDP-glucose 4,6-dehydratase rmlB Rv3464 TR:O06329 (EMBL:Z95390) (331 aa); FastA score E(): 0, 84.0% identity in 331 aa overlap, and TR:Q50556 (EMBL:U43540) (329 aa); FastA score E(): 0, 78.2% identity in 331 aa overlap, and to others e.g. Streptococcus pneumoniae dTDP-glucose-4,6-dehydratase cpsN TR:O54611 (EMBL:AF030364) (349 aa); FastA score E(): 0, 59.9% identity in 334 aa overlap. Contains Pfam match to entry PF01370 Epimerase, NAD dependent epimerase/dehydratase family. Similar to ML0204, ML0751, ML1942 and ML2428"
 /codon_start=1
 /transl_table=11
 /product="dTPP-glucose 4,6-dehydratase"
 /protein_id="CAC30919.1"
 /db_xref="GI:13093619"
 /db_xref="GOA:Q9CBH7"
 /db_xref="TrEMBL:Q9CBH7"
 /transl_table=11
 /product="PPE family protein (pseudogene)"
 complement (5464..5649)
 /gene="ML1968"
 /pseudo
 complement (5464..5649)

FEATURES

source

gene

CDS

gene

CDS

misc_feature

gene

CDS

misc_feature

gene

CDS

misc_feature

gene

CDS

gene

CDS

DHNSAVMOILEKGQIGRTYLICAGHNHNLTVLRTLQMMGRDNDPFDHVTDRVGH
 LRYADPTPLYNELCWAPHANFDEGLRATIDWYRNESWWRPLKDAVEARVEGR"
 495..1424
 /gene="rmlB"
 /note="Pfam match to entry PF01370 Epimerase, NAD dependent epimerase/dehydratase family, score 470.30, E-value 1.6e-137"
 1486..2094
 /gene="rmlC"
 /note="synonym: ML1965"
 1486..2094
 /gene="rmlC"
 /note="Similar to M. tuberculosis dTDP-4-dehydrorhamnose 3,5-epimerase rmlC Rv3465 TR:O06330 (EMBL:Z95390) (202 aa); FastA score E(): 0, 75.4% identity in 199 aa overlap, and to many others e.g. Streptomyces griseus dTDP-4-dehydrorhamnose 3,5-epimerase SW:STRM_STRGR (P29783) (200 aa); FastA score E(): 0, 45.6% identity in 193 aa overlap. Contains Pfam match to entry PF00908 dTDP_sugar_isom, dTDP-4-dehydrorhamnose 3,5-epimerase." /codon_start=1
 /transl_table=11
 /product="dTDP-4-dehydrorhamnose 3,5-epimerase"
 /protein_id="CAC30920.1"
 /db_xref="GI:13093620"
 /db_xref="GOA:Q9X7A4"
 /db_xref="TrEMBL:Q9X7A4"
 /translation="MNVELDISGAIWEITPTIHSRGMFFEWLTASRFQAFTHRLD VRQANCSVSSAGVLGRHFAQPPSQAKYTCVTGVSFVDVDIRLGSPTFGWTSIT LNSNRKTIYISLHAGFLALQDNSTWYLCAGYNPAREHAICATDPLDLADWPV DGVDTLHLSKRDATAPNLDIRASGLLPTWDETQNFANLGSK"
 1492..2025
 /gene="rmlC"
 /note="Pfam match to entry PF00908 dTDP_sugar_isom, dTDP-4-dehydrorhamnose 3,5-epimerase, score 308.80, E-value 6.5e-89"
 2139..2624
 /gene="lpqH"
 /note="synonym: ML1966"
 2139..2624
 /gene="lpqH"
 /note="Similar to M. tuberculosis 19 kDa lipoprotein antigen precursor lpqH Rv3763. Contains a possible N-terminal signal sequence and a PS00013 Prokaryotic membrane lipoprotein lipid attachment site." /codon_start=1
 /transl_table=11
 /product="possible lipoprotein"
 /protein_id="CAC30921.1"
 /db_xref="GI:13093621"
 /db_xref="TrEMBL:Q9X7A5"
 /translation="MRHKLIAIYAVTIMAGAGCGGTQAPTPSVSKTNSPTTVA SSIPDAAGSTKVTIGSQPKVSGPVVCSCTTGKFSIAGDMITGVUGLEPDASVW NAGLGTIDGVVIAFTGVSENNATKNTVIQGTASGVNDTGGQIHKFSFEIVC R"
 2169..2201
 /gene="lpqH"
 /note="PS00013 Prokaryotic membrane lipoprotein lipid attachment site"
 complement (4768..4995)
 /gene="ML1967"
 /pseudo
 complement (4768..4995)
 /gene="ML1967"
 /note="Possible pseudogene of M. tuberculosis paralog pPE"
 /pseudo
 /codon_start=1
 /transl_table=11
 /product="PPE family protein (pseudogene)"
 complement (5464..5649)
 /gene="ML1968"
 /pseudo
 complement (5464..5649)

```

/gene="ML1969"
/Note="Possible pseudogene of M. tuberculosis paralog PE"
/pseudo
/codon_start=1
/transl_table=11
/product="PE family protein (pseudogene)"
/complement(5661..8043)
/Note="Dispersed repeat, LEPREP, copy 6"
/complement(5848..5946)
/gene="ML1969"
/pseudo
/complement(5848..5946)
/gene="ML1969"
/Note="Similar to Agrobacterium tumefaciens transposase
TR:Q44454 (EMBL:Z18270) (366 aa) fasta scores: E():
0.0002, 51.5% id in 33 aa, and to Pseudomonas putida
transposase tnpal TR:Q9R9U9 (EMBL:AJ245436) (355 aa) fasta
scores: E(): 0.00088, 48.5% id in 33 aa"
/pseudo
/codon_start=1
/transl_table=11
/complement(6211..7354)
/gene="ML1970"
/pseudo
/complement(6211..7354)
/gene="ML1970"
/Note="Similar to many e.g. Cryphonectria parasitica
(Chestnut blight fungus) putative maturase TR:AAF27656
(EMBL:AF218567) (778 aa) fasta scores: E(): 7.8e-11, 29.2%
id in 216 aa"
/pseudo
/codon_start=1
/transl_table=11
/product="putative group II intron maturase-related
protein"
/complement(6674..6727)
/gene="ML1970"
/complement(7559..7698)
/gene="ML1971"
/pseudo
/complement(7559..7698)
/gene="ML1971"
/Note="Similar to Pseudomonas putida transposase tnpal
TR:Q9R9U9 (EMBL:AJ245436) (355 aa) fasta scores: E():
3.4e-05, 47.8% id in 46 aa, and to Agrobacterium
tumefaciens transposase TR:Q44454 (EMBL:Z18270) (366 aa)
fasta scores: E(): 0.00026, 41.3% id in 46 aa"

Query Match      91.1%; Score 16.4; DB 1; Length 342300;
Best Local Similarity 94.4%; Pred. No. 5e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GAGAACACACAGATTCG 18
Db 110620 GAGTACACACAGATTCG 110637

RESULT 13
AX481359
LOCUS AX481359 19 bp DNA linear PAT 16-AUG-2002
DEFINITION Sequence 6 from Patent EP1225232.
ACCESSION AX481359
VERSION AX481359.1 GI:22316280
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
REFERENCE
AUTHORS Rubin,B.Y. and Anderson,S.L.
TITLE Detection of mutations in a gene encoding lkappab
kinase-complex-associated protein to diagnose familial dysautonomia
JOURNAL Patent: EP 1225232-A 6 24-JUL-2002;

```

```

FEATURES
Source
1..19
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match      88.9%; Score 16; DB 6; Length 19;
Best Local Similarity 100.0%; Pred. No. 1.1e+03;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAGAACACACAGATTC 16
Db 1 GAGAACACACAGATTC 16

RESULT 14
AC009112
LOCUS AC009112 64167 bp DNA linear HTG 26-JAN-2000
DEFINITION Homo sapiens chromosome 16 clone RP11-468p11, LOW-PASS SEQUENCE
SAMPLING.
ACCESSION AC009112 GI:6758904
VERSION AC009112.2 GI:6758904
KEYWORDS HTG; HTGS PHASEO.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 64167)
REFERENCE Sequencing of Human Chromosome 16
AUTHORS Unpublished
TITLE DOE Joint Genome Institute.
REFERENCE 2 (bases 1 to 64167)
AUTHORS DOE Joint Genome Institute.
TITLE Direct Submission
JOURNAL Submitted (03-AUG-1999) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
On Jan 26, 2000 this sequence version replaced gi:5685963.
-----Genome Center
Center: Joint Genome Institute
Center Code: JGI
Web site: http://www.jgi.doe.gov

-----
* NOTE: This record contains 82 individual
* sequencing reads that have not been assembled into
* contigs. Runs of N are used to separate the reads
* and the order in which they appear is completely
* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene-rich and allows
* overlap relationships among clones to be deduced
* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.
* 1 726: contig of 726 bp in length
* gap of unknown length
* 727 1457: contig of 731 bp in length
* gap of unknown length
* 1458 2142: contig of 685 bp in length
* gap of unknown length
* 2143 2680: contig of 538 bp in length
* gap of unknown length
* 2681 3386: contig of 706 bp in length
* gap of unknown length
* 3387 4102: contig of 716 bp in length
* gap of unknown length
* 4103 4654: contig of 552 bp in length
* gap of unknown length
* 4655 5221: contig of 567 bp in length
* gap of unknown length
* 5222 6170: contig of 949 bp in length
* gap of unknown length

```

* 6171	6805:	contig of 635 bp in length	* 31951	gap of unknown length
* 6806	7053:	gap of unknown length	* 32650:	contig of 700 bp in length
* 7054	7674:	contig of 248 bp in length	* 33308:	gap of unknown length
* 7675	8411:	gap of unknown length	* 33965:	contig of 657 bp in length
* 8412	9038:	gap of unknown length	* 34708:	gap of unknown length
* 9039	9748:	contig of 627 bp in length	* 35781:	contig of 743 bp in length
* 9749	10370:	gap of unknown length	* 36274:	gap of unknown length
* 10371	11089:	contig of 622 bp in length	* 37097:	contig of 1073 bp in length
* 11090	11781:	gap of unknown length	* 37759:	contig of 493 bp in length
* 11782	12496:	contig of 719 bp in length	* 38327:	gap of unknown length
* 12497	13259:	gap of unknown length	* 39523:	contig of 568 bp in length
* 13260	13372:	contig of 692 bp in length	* 40180:	gap of unknown length
* 13373	14620:	contig of 715 bp in length	* 40800:	gap of unknown length
* 14621	15606:	gap of unknown length	* 41388:	contig of 1196 bp in length
* 15607	16393:	contig of 1248 bp in length	* 42391:	gap of unknown length
* 16394	17093:	gap of unknown length	* 42985:	contig of 594 bp in length
* 17094	17798:	gap of unknown length	* 43747:	gap of unknown length
* 17799	18473:	contig of 705 bp in length	* 44488:	contig of 742 bp in length
* 18474	19192:	gap of unknown length	* 45659:	gap of unknown length
* 19193	19908:	contig of 675 bp in length	* 45791:	contig of 131 bp in length
* 19907	20671:	gap of unknown length	* 46221:	contig of 431 bp in length
* 20672	21341:	gap of unknown length	* 47164:	gap of unknown length
* 21342	22132:	contig of 765 bp in length	* 47165:	contig of 943 bp in length
* 22133	22791:	gap of unknown length	* 48194:	gap of unknown length
* 22792	23584:	gap of unknown length	* 48195:	contig of 1030 bp in length
* 23555	24087:	contig of 670 bp in length	* 49050:	gap of unknown length
* 24088	24798:	gap of unknown length	* 49702:	contig of 856 bp in length
* 24799	25531:	contig of 533 bp in length	* 49883:	gap of unknown length
* 25532	25668:	gap of unknown length	* 49884:	contig of 181 bp in length
* 25669	26491:	gap of unknown length	* 51039:	gap of unknown length
* 26492	27242:	contig of 733 bp in length	* 52165:	contig of 1156 bp in length
* 27243	27842:	gap of unknown length	* 53296:	gap of unknown length
* 27843	28475:	contig of 137 bp in length	* 54764:	contig of 1131 bp in length
* 28476	29146:	gap of unknown length	* 55985:	gap of unknown length
* 29147	29734:	contig of 823 bp in length	* 57280:	contig of 1221 bp in length
* 29735	31298:	gap of unknown length	* 58306:	gap of unknown length
* 31299	31950:	contig of 751 bp in length	* 59481:	contig of 1026 bp in length
		gap of unknown length	* 60351:	gap of unknown length
		contig of 600 bp in length	* 63313:	contig of 2962 bp in length
		gap of unknown length	* 64167:	gap of unknown length
		contig of 633 bp in length	* 63314	Location/Qualifiers
		gap of unknown length		
		contig of 671 bp in length		
		gap of unknown length		
		contig of 588 bp in length		
		gap of unknown length		
		contig of 1564 bp in length		
		gap of unknown length		
		contig of 652 bp in length		

Query Match 88.9%; Score 16; DB 2; Length 64167;
 Best Local Similarity 100.0%; Pred. No. 8.9e+02;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAGAACACACAGATTC 16
 DB 21230 GAGAACACACAGATTC 21245
 |||||

RESULT 15
 LOCUS AX676048 66479 bp DNA linear PAT 27-MAR-2003
 DEFINITION Sequence 1 from Patent WO02059381.
 ACCESSION AX676048
 VERSION AX676048.1 GI:29333739
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE 1
 AUTHORS Slangenaupt, S. and Guehella, J.F.
 TITLE Gene for identifying individuals with familial dysautonomia
 JOURNAL Patent: WO 02059381-A 1 01-AUG-2002;
 The General Hospital Corporation (US)

FEATURES
 source
 1. 66479
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"

ORIGIN
 Query Match 88.9%; Score 16; DB 6; Length 66479;
 Best Local Similarity 100.0%; Pred. No. 8.9e+02;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAGAACACACAGATTC 16
 DB 33979 GAGAACACACAGATTC 33994
 |||||

RESULT 16
 LOCUS AL359692/c 78376 bp DNA linear PRI 18-JUL-2001
 DEFINITION Human DNA sequence from clone RP11-3J11 on chromosome 9, complete
 sequence.

ACCESSION AL359692
 VERSION AL359692.9 GI:14970800
 KEYWORDS HTG.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE 1
 AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 TITLE Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 JOURNAL Sycamore, N.
 Direct Submission

COMMENT
 Submitted (18-JUL-2001) Sanger Centre, Hinxton, Cambridgeshire,
 CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
 requests: clonerequests@sanger.ac.uk
 On Jul 19, 2001 this sequence version replaced gi:14715342.
 During sequence assembly data is compared from overlapping clones.
 Where differences are found these are annotated as variations
 together with a note of the overlapping clone name. Note that the
 variation annotation may not be found in the sequence submission
 corresponding to the overlapping clone, as we submit sequences with
 only a small overlap as described above.
 This sequence was finished as follows unless otherwise noted: all
 regions were either double-stranded or sequenced with an alternate
 chemistry or covered by high quality data (i.e., phred quality >= 30);
 an attempt was made to resolve all sequencing problems, such
 as compressions and repeats; all regions were covered by at least
 one plasmid subclone or more than one M13 subclone; and the

assembly was confirmed by restriction digest. The following
 abbreviations are used to associate primary accession numbers given
 in the feature table with their source databases: Em: EMBL; Sw:,
 SWISSPROT; Tr: TREMBL; Wp:, WORMPEP; Information on the WORMPEP
 database can be found at
 http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
 was generated from part of bacterial clone contigs of human
 chromosome 9, constructed by the Sanger Centre Chromosome 9 Mapping
 Group. Further information can be found at
 http://www.sanger.ac.uk/HGP/Chr9
 RP11-3J11 is from the library RPCI-11.1 constructed by the group of
 Pieter de Jong. For further details see
 http://www.chori.org/bacpac/home.htm
 VECTOR: pBAC3.6

IMPORTANT: This sequence is not the entire insert of clone
 RP11-3J11 it may be shorter because we sequence overlapping
 sections only once, except for a 100 base overlap.
 The true left end of clone RP11-3J11 is at 1 in this sequence. The
 true left end of clone RP11-11522 is at 76377 in this sequence.
 The true right end of clone RP11-339N8 is at 76382 in this
 sequence.

FEATURES Location/Qualifiers

1. 78376
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="9"
 /clone="RP11-3J11"
 /clone_lib="RPCI-11.1"
 7. 829
 /note="LIPB1 repeat: matches 5343. .6155 of consensus"
 991. 1328
 /note="MER46B repeat: matches 3. .231 of consensus"
 2199. 2234
 /note="18 copies 2 mer tt 86% conserved"
 3191. 3349
 /note="AluSg/x repeat: matches 153. .311 of consensus"
 4250. 4537
 /note="AluSx repeat: matches 1. .288 of consensus"
 5314. 5579
 /note="AluJb repeat: matches 1. .299 of consensus"
 5693. 5863
 /note="L1MC4 repeat: matches 6481. .6679 of consensus"
 5881. 6013
 /note="AluJb repeat: matches 1. .135 of consensus"
 6533. 6644
 /note="MIR repeat: matches 92. .211 of consensus"
 7645. 7979
 /note="MT1J repeat: matches 3. .363 of consensus"
 9165. 9774
 /note="MER21B repeat: matches 179. .787 of consensus"
 9775. 9848
 /note="MER84 repeat: matches 1. .76 of consensus"
 9849. 10232
 /note="MER84 repeat: matches 3. .373 of consensus"
 10233. 10535
 /note="AluSx repeat: matches 1. .302 of consensus"
 10536. 10669
 /note="MER84 repeat: matches 373. .508 of consensus"
 10669. 10846
 /note="MER21B repeat: matches 7. .185 of consensus"
 11168. 11901
 /note="LIPAL4 repeat: matches 5403. .6149 of consensus"
 12172. 12468
 /note="AluSg repeat: matches 1. .297 of consensus"
 13074. 13362
 /note="AluSx repeat: matches 3. .294 of consensus"
 13411. 13533
 /note="MIR repeat: matches 117. .239 of consensus"
 13616. 13669
 /note="L2 repeat: matches 2445. .2498 of consensus"
 13733. 13841
 /note="MER53 repeat: matches 89. .188 of consensus"

```

repeat_region 13996. .14113
/note="AluSg/x repeat: matches 178. .295 of consensus"
repeat_region 14769. .15142
/note="MSTA repeat: matches 1. .423 of consensus"
repeat_region 15274. .15333
/note="MERA repeat: matches 14. .74 of consensus"
repeat_region 15480. .15600
/note="MERSB repeat: matches 54. .175 of consensus"
repeat_region 15640. .15689
/note="MSTA repeat: matches 1. .49 of consensus"
repeat_region 15861. .16180
/note="AluB repeat: matches 1. .302 of consensus"
repeat_region 17189. .17296
/note="MIR repeat: matches 100. .206 of consensus"
repeat_region 17614. .17738
/note="MERSB repeat: matches 54. .174 of consensus"
repeat_region 17769. .18234
/note="L1MD1 repeat: matches 5748. .6223 of consensus"
repeat_region 18235. .18529
/note="AluJ repeat: matches 1. .292 of consensus"
repeat_region 18530. .19071
/note="L1MD1 repeat: matches 5238. .5748 of consensus"
repeat_region 19132. .19432
/note="AluSp repeat: matches 1. .313 of consensus"
repeat_region 19441. .19472
/note="16 copies 2 mer tt 84% conserved"
repeat_region 19498. .19859
/note="THEIA repeat: matches 1. .353 of consensus"
repeat_region 19860. .21021
/note="THEIA-internal repeat: matches 427. .1580 of consensus"
repeat_region 21022. .21314
/note="AluSg repeat: matches 1. .302 of consensus"
repeat_region 21315. .21711
/note="THEIA-internal repeat: matches 29. .427 of consensus"
repeat_region 21718. .21818
/note="THEIB repeat: matches 265. .364 of consensus"
repeat_region 21813. .21845
/note="Alu repeat: matches 3. .35 of consensus"
repeat_region 21878. .22184
/note="AluSg repeat: matches 1. .311 of consensus"
repeat_region 22578. .22715
/note="MIR repeat: matches 74. .210 of consensus"
repeat_region 22693. .22848
/note="L2 repeat: matches 2572. .2745 of consensus"
repeat_region 24368. .24434
/note="MIR repeat: matches 71. .137 of consensus"
repeat_region 25259. .25338
/note="L2 repeat: matches 2605. .2695 of consensus"
repeat_region 25474. .25775
/note="AluSx repeat: matches 1. .301 of consensus"
misc_feature 26871. .27694
/note="CpG island"
/evidence=not_experimental
repeat_region 28479. .28806
/note="L2 repeat: matches 2157. .2496 of consensus"
repeat_region 28807. .29114
/note="AluSg1 repeat: matches 1. .305 of consensus"
repeat_region 29115. .29236
/note="L2 repeat: matches 1963. .2157 of consensus"
repeat_region 29748. .30050
/note="AluSg repeat: matches 1. .303 of consensus"
repeat_region 30475. .30548
/note="MIR repeat: matches 33. .107 of consensus"
repeat_region 30654. .30701
/note="24 copies 2 mer ac 95% conserved"
repeat_region 31870. .32343
/note="MTRD repeat: matches 16. .505 of consensus"
repeat_region 33199. .33414
/note="L2 repeat: matches 2116. .2353 of consensus"
repeat_region 33415. .33568
/note="MERSB repeat: matches 186. .340 of consensus"

```

```

repeat_region 33658. .33918
/note="L2 repeat: matches 2403. .2710 of consensus"
repeat_region 37241. .37362
/note="Charlie4 repeat: matches 1823. .1956 of consensus"
repeat_region 37669. .37895
/note="MER20 repeat: matches 1. .218 of consensus"
repeat_region 38589. .38875
/note="AluSg repeat: matches 1. .310 of consensus"
repeat_region 39767. .39947
/note="MSTB repeat: matches 1. .183 of consensus"
misc_feature 40006. .40255
/note="match: STS: Em:G14836"
misc_feature 40477. .40798
/note="match: STS: Em:G21606"
repeat_region 41261. .41395
/note="L1M2 repeat: matches 6041. .6161 of consensus"
repeat_region 41739. .41805
/note="TIGER2 repeat: matches 2653. .2718 of consensus"
repeat_region 41806. .42093
/note="AluB repeat: matches 1. .300 of consensus"
repeat_region 42094. .43429
/note="TIGER2 repeat: matches 1302. .2653 of consensus"
repeat_region 43429. .44232
/note="TIGER2 repeat: matches 297. .1113 of consensus"
repeat_region 44353. .44664

Query Match 88.9%; Score 16; DB 9; Length 78376;
Best Local Similarity 100.0%; Pred. No. 8.8e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAGAACACACAGATTC 16
|||||
Db 71868 GAGAACACACAGATTC 71853

RESULT 17
AC135161/c
LOCUS AC135161 88701 bp DNA linear PLN 15-MAR-2004
DEFINITION Medicago truncatula clone mth2-30k24, complete sequence.
ACCESSION AC135161
VERSION AC135161.11 GI:45434547
KEYWORDS HTG.
SOURCE Medicago truncatula (barrel medic)
ORGANISM Medicago truncatula
Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
Medicago.
REFERENCE 1 (bases 1 to 88701)
AUTHORS Shaull,S., Lin,S., Dixon,R., May,G., Sumner,L., Gonzales,B.,
Cook,D., Kim,D. and Roe,B.A.
Medicago truncatula BAC Clone mth2-30k24
Unpublished
REFERENCE 2 (bases 1 to 88701)
AUTHORS Shaull,S., Lin,S., Dixon,R., May,G., Sumner,L., Gonzales,B.,
Cook,D., Kim,D. and Roe,B.A.
Direct Submission
JOURNAL Submitted (08-OCT-2002) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
REFERENCE 3 (bases 1 to 88701)
AUTHORS Shaull,S., Lin,S., Dixon,R., May,G., Sumner,L., Gonzales,B.,
Cook,D., Kim,D. and Roe,B.A.
Direct Submission
JOURNAL Submitted (06-MAY-2003) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
REFERENCE 4 (bases 1 to 88701)
AUTHORS Shaull,S., Lin,S., Dixon,R., May,G., Sumner,L., Gonzales,B.,
Cook,D., Kim,D. and Roe,B.A.
Direct Submission
JOURNAL Submitted (21-MAY-2003) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,

```

OK 73019, USA
5 (bases 1 to 88701)
Shauli, S., Lin, S., Dixon, R., May, G., Sumner, L., Gonzales, B.,
Cook, D., Kim, D. and Roe, B.A.
Direct Submission
Submitted (12-MAR-2004) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
6 (bases 1 to 88701)
Shauli, S., Lin, S., Dixon, R., May, G., Sumner, L., Gonzales, B.,
Cook, D., Kim, D. and Roe, B.A.
Direct Submission
Submitted (15-MAR-2004) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
On Mar 15, 2004 this sequence version replaced gi:45384562.
----- Genome Center
Center: Department Of Chemistry And Biochemistry
The University Of Oklahoma
Center code: UOKNOR

FEATURES
source Location/Qualifiers
1..88701
/organism="Medicago truncatula"
/mol_type="genomic DNA"
/db_xref="taxon:3880"
/clone="mth2-30k24"
/clone_lib="Medicago truncatula BAC library H2"
ORIGIN
Query Match 88.9%; Score 16; DB 8; Length 88701;
Best Local Similarity 100.0%; Pred. No. 8.9e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 AGAACACACAGATTTCG 17
Db 23358 AGAACACACAGATTTCG 23343
RESULT 18
CR382130_35/c
WPCOMMENT
Sequence split into 37 fragments LOCUS CR382130 Accession CR382130
Fragment Name Begin End
CR382130_00 1 110000
CR382130_01 100001 210000
CR382130_02 200001 310000
CR382130_03 300001 410000
CR382130_04 400001 510000
CR382130_05 500001 610000
CR382130_06 600001 710000
CR382130_07 700001 810000
CR382130_08 800001 910000
CR382130_09 900001 1010000
CR382130_10 1000001 1110000
CR382130_11 1100001 1210000
CR382130_12 1200001 1249565
Continuation (7 of 13) of CR382133 from base 600001 (CR382133 Debaryomyces hansenii chron

CR382130_29 2900001 3010000
CR382130_30 3000001 3110000
CR382130_31 3100001 3210000
CR382130_32 3200001 3310000
CR382130_33 3300001 3410000
CR382130_34 3400001 3510000
CR382130_35 3500001 3610000
CR382130_36 3600001 3632272
Continuation (36 of 37) of CR382130 from base 3500001 (CR382130 Yarrowia lipolytica chron
Query Match 88.9%; Score 16; DB 8; Length 110000;
Best Local Similarity 100.0%; Pred. No. 8.7e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GAGAACACACAGATTTC 16
Db 77927 GAGAACACACAGATTTC 77912
RESULT 19
CR382133_06/c
WPCOMMENT
Sequence split into 13 fragments LOCUS CR382133 Accession CR382133
Fragment Name Begin End
CR382133_00 1 110000
CR382133_01 100001 210000
CR382133_02 200001 310000
CR382133_03 300001 410000
CR382133_04 400001 510000
CR382133_05 500001 610000
CR382133_06 600001 710000
CR382133_07 700001 810000
CR382133_08 800001 910000
CR382133_09 900001 1010000
CR382133_10 1000001 1110000
CR382133_11 1100001 1210000
CR382133_12 1200001 1249565
Continuation (7 of 13) of CR382133 from base 600001 (CR382133 Debaryomyces hansenii chron
Query Match 88.9%; Score 16; DB 8; Length 110000;
Best Local Similarity 100.0%; Pred. No. 8.7e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 AGAACACACAGATTTCG 17
Db 17211 AGAACACACAGATTTCG 17196
RESULT 20
AC011700/c
LOCUS
DEFINITION Homo sapiens Chromosome 1 BAC RP11-478J18 (Roswell Park Cancer
Institute Human BAC Library) complete sequence.
ACCESSION AC011700
VERSION AC011700.5 GI:8439783
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 154732)
Muzny, D.M., Adams, C., Bailey, M., Barberia, J., Blankenburg, K.,
Bodota, B., Buck, J., Bowie, S., Brooks, A., Bunay, C., Bunac, C.,
Burkett, C., Burrows, J., Carter, M., Chacko, J., Chen, Z., Cox, C.,
David, R., Delgado, O., Deshazo, D., Ding, Y., Domah-Rashid, N.,
Dugan-Rocha, S., Durbin, K.J., Fernandez, C., Ferraguto, D.,
Forcum-Tansey, J., Frantz, P., Ganesh, R., Gorrell, J.H., Gorrell, L.L.,
Guevara, W., Harris, K., Hernandez, J., Hodgson, A., Hoques, M.,
Holloway, C., Hosak, H., Issar, A., Jackson, L.E., Jackson, L., Jia, Y.,
Jones, M., Kelly, S., Kneitz, S., Kondejewski, N., Kong, Y., Kovar, C.,
Lau, S., Leal, B., Lee, E., Li, Z., Lichtarge, O., Liu, J., Liu, W.,
Logan, O., Lu, J., Lucier, R., Marondel, I., Martin, R., Martinez, C.,
McLeod, M.P., Mei, G., Merscher, S., Miller, A., Montgomery, K.T.,


```

repeat_region      26242..26300
                    /rpt_family="A-rich"
repeat_region      26918..26970
                    /rpt_family="(TA)n"
repeat_region      31000..31061
                    /rpt_family="(CA)n"
repeat_region      31068..31090
                    /rpt_family="(TC)n"
repeat_region      31548..31587
                    /rpt_family="(T)n"
misc_feature       31825..32144
                    /note="Region: H.sapiens CpG island DNA genomic MseI
                    fragment, clone 33a11, Z60720"
misc_feature       32403..32608
                    /note="Region: H.sapiens CpG island DNA genomic MseI
                    fragment, clone 33a11, Z60719"
repeat_region      32655..33310
                    /rpt_family="(TC)n"
repeat_region      33447..33626
                    /rpt_family="GA-rich"
misc_feature       33899..34029
                    /function="Low coverage"
repeat_region      33973..34018
                    /rpt_family="(CCCCCA)n"
repeat_region      34145..34172
                    /rpt_family="(T)n"
repeat_region      36959..37011
                    /rpt_family="(CCA)n"
misc_feature       37522..37889
                    /note="Region: ox03f03.s1 Homo sapiens cDNA, AI033409"
repeat_region      complement(39640..39663)

Query Match      88.9%; Score 16; DB 9; Length 154732;
Best Local Similarity 100.0%; Pred. No. 8.6e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAGACACACAGATTTC 16
Db 87079 GAGACACACAGATTTC 87064
|||||
|||||

RESULT 21
AC129818/c
LOCUS              157381 bp      DNA      linear      HTG 20-NOV-2002
DEFINITION         Rattus norvegicus clone CH230-392M8, WORKING DRAFT SEQUENCE, 3
UNORDERED PIECES.
ACCESSION           AC129818
VERSION             GI:25137966
KEYWORDS            HTG; HTGS PHASE1; HTGS DRAFT; HTGS_FULLTOP.
SOURCE              Rattus norvegicus
ORGANISM            Rattus norvegicus
                    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                    Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
                    Rattus.
REFERENCE           1 (bases 1 to 157381)
AUTHORS            Muzny,D,Marle,E, Metzker,M, Lee,A, Abramzon,S, Adams,C, Alder,J,
                    Allen,C, Allen,H, Alsbrooks,S, Amin,A, Anguiano,D,
                    Anyalebechi,V, Aoyagi,A, Ayodeji,M, Baca,E, Baden,H,
                    Baldwin,D, Bandaranaike,D, Barber,M, Barnstead,M, Benahmed,F,
                    Biswal,K, Blair,J, Blankenburg,K, Blyth,P, Brown,M,
                    Bryant,N, Buhay,C, Burch,P, Burrell,K, Calderon,E,
                    Cardenas,V, Carter,K, Cavazos,I, Cesar,H, Center,A,
                    Chacko,J, Chavez,D, Chen,G, Chen,R, Chen,Y, Chen,Z, Chu,J,
                    Cleveland,C, Cockrell,R, Cox,C, Coyle,M, Cree,A, D'Souza,L,
                    Davila,M,L, Davis,C, Davy-Carroll,L, De Anda,C, Dederich,D,
                    Delgado,O, Denson,S, Deramo,C, Ding,Y, Dinh,H, Divya,K,
                    Draper,H, Dugan-Rocha,S, Dunn,A, Durbin,K, Duval,B, Evans,K,
                    Egan,A, Escotto,M, Eugene,C, Evans,C,A, Falls,T, Fan,G,
                    Fernandez,S, Finley,M, Flagg,N, Forbes,L, Foster,M, Foster,P,
                    Fraser,C,M, Gabis,A, Ganta,R, Garcia,A, Garner,T, Garza,M,
                    Gebregorgis,E, Geer,K, Gill,R, Grady,M, Guerra,W, Guevara,W,
                    Gunaratne,P, Haaland,W, Hamil,C, Hamilton,C, Hamilton,K,
                    Harvey,Y, Havlak,P, Hawes,A, Henderson,N, Hernandez,J,
                    Muzny,D,Marle,E, Metzker,M, Lee,A, Abramzon,S, Adams,C, Alder,J,
                    Allen,C, Allen,H, Alsbrooks,S, Amin,A, Anguiano,D,
                    Anyalebechi,V, Aoyagi,A, Ayodeji,M, Baca,E, Baden,H,
                    Baldwin,D, Bandaranaike,D, Barber,M, Barnstead,M, Benahmed,F,
                    Biswal,K, Blair,J, Blankenburg,K, Blyth,P, Brown,M,
                    Bryant,N, Buhay,C, Burch,P, Burrell,K, Calderon,E,
                    Cardenas,V, Carter,K, Cavazos,I, Cesar,H, Center,A,
                    Chacko,J, Chavez,D, Chen,G, Chen,R, Chen,Y, Chen,Z, Chu,J,
                    Cleveland,C, Cockrell,R, Cox,C, Coyle,M, Cree,A, D'Souza,L,
                    Davila,M,L, Davis,C, Davy-Carroll,L, De Anda,C, Dederich,D,
                    Delgado,O, Denson,S, Deramo,C, Ding,Y, Dinh,H, Divya,K,
                    Draper,H, Dugan-Rocha,S, Dunn,A, Durbin,K, Duval,B, Evans,K,
                    Egan,A, Escotto,M, Eugene,C, Evans,C,A, Falls,T, Fan,G,
                    Fernandez,S, Finley,M, Flagg,N, Forbes,L, Foster,M, Foster,P,
                    Fraser,C,M, Gabis,A, Ganta,R, Garcia,A, Garner,T, Garza,M,
                    Gebregorgis,E, Geer,K, Gill,R, Grady,M, Guerra,W, Guevara,W,
                    Gunaratne,P, Haaland,W, Hamil,C, Hamilton,C, Hamilton,K,
                    Harvey,Y, Havlak,P, Hawes,A, Henderson,N, Hernandez,J,

```

```

Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hoques,M.,
Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A.,
Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,
Karpathy,S., Kelly,S., Khan,L., Khan,Z., King,L., Kovar,C.,
Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,
Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
Lorensuhera,L., Loulseged,H., Lozado,R.J., Lu,X., Ma,J.,
Maheshwari,M., Mahindaratne,M., Mahmoud,M., Malloy,K., Mangum,A.,
Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E.,
Mawhney,S., McLeod,M.P., McNeill,T.Z., Meenen,E.,
Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S.,
Morgan,M., Morris,K., Morris,S., Muridasa,M., Murphy,M., Nair,L.,
Narkvis,C., Neal,D., Newton,N., Nguyen,N., Norris,S.,
Nwaokwelenh,O., Okwuonu,G., Olarnpusagcon,A., Pal,S., Parks,K.,
Pasternak,S., Paul,H., Perez,A., Popovic,D., Primus,E., Pu,L.,
Puazo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R.,
Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F.,
Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J.,
Sanders,W., Savery,G., Scherer,S., Scott,G., Shatsman,S., Shen,H.,
Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Sosa,J.,
Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J.,
Steimle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C.,
Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K.,
Valas,R., Vera,V., Villasana,D., Waldron,L., Walker,B., Wang,J.,
Wang,O., Wang,S., Warren,J., Warren,R., Wei,X., White,F.,
Williams,G., Willson,R., Wleczky,R., Wooden,H., Worley,K.,
Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V.,
Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von
Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O.,
Weinstock,G. and Gibbs,R.A.
Direct Submission
2 (bases 1 to 157381)
Unpublished
2 (bases 1 to 157381)
Worley,K.C.
Direct Submission
Submitted (03-AUG-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 157381)
Rat Genome Sequencing Consortium.
Direct Submission
Submitted (20-NOV-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
The sequence in this assembly is a combination of BAC based reads
and whole genome shotgun sequencing reads assembled using Atlas
(http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
in the feature table below represents a scaffold in the Atlas
assembly (a 'contig-scaffold'). Within each contig-scaffold,
individual sequence contigs are ordered and oriented, and separated
by sized gaps filled with Ns to the estimated size. The sequence
may extend beyond the ends of the clone and there may be sequence
contigs within a contig-scaffold that consist entirely of whole
genome shotgun sequence reads. Both end sequences and whole genome
shotgun sequence only contigs will be indicated in the feature
table.
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GWTE
Center clone name: CH230-392M8
----- Summary Statistics
Assembly program: Phrap; version 0.990329
Consensus quality: 153926 bases at least Q40
Consensus quality: 154705 bases at least Q30
Consensus quality: 155316 bases at least Q20
Estimated insert size: 156048; sum-of-contigs estimation
Quality coverage: 10x in Q20 bases; sum-of-contigs estimation

```

***** NOTE: Estimated insert size may differ from sequence length
 (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
 * NOTE: This is a "working draft" sequence. It currently
 * consists of 3 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

1 55243: contig of 55243 bp in length
 55244 55343: gap of unknown length
 55344 153147: contig of 97804 bp in length
 153148 153247: gap of unknown length
 153248 157381: contig of 4134 bp in length.

FEATURES
 source
 1. 157381
 Location/Qualifiers
 /organism="Rattus norvegicus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10116"
 /clone="CH230-392M8"

misc_feature
 1. 1061
 /note="wgs end_extension
 clone_end:Sp6"

misc_feature
 1117..1750
 /note="clone boundary
 clone_end:Sp6"

misc_feature
 5207..5765
 /note="clone boundary
 clone_end:T7
 site:
 end_sequence:BZ276820"
 5207..5765
 /note="clone boundary
 clone_end:T7
 site:
 end_sequence:BZ276819"

ORIGIN

Query Match 88.9%; Score 16; DB 2; Length 157381;
 Best Local Similarity 100.0%; Pred. No. 8.6e+02;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAGAACACACAGATTC 16
 |||||
 Db 4779 GAGAACACACAGATTC 4764

RESULT 22
 AC007341 168612 bp DNA linear PRI 27-MAR-2003
 LOCUS Homo sapiens chromosome 16 clone RP11-381L15, complete sequence.
 DEFINITION
 AC007341
 VERSION AC007341.10 GI:29294011
 KEYWORDS HTG.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE
 AUTHORS Jones, D., Munt, M., Doggett, N., Munk, C., Saunders, E., Robinson, D.,
 Bryant, J., Tesmer, J., Chasteen, L., Thompson, S., Goodwin, L.,
 Campbell, C., Fawcett, J., Meincke, L., Longmire, J., White, S., Tatum, O.,
 Campbell, C., Fawcett, J., Maltbie, M., Bussod, M., Sutherland, R.,
 McMurtry, K., Han, C. and Deaven, L.
 TITLE Direct Submission
 JOURNAL Submitted (21-APR-1999) Center for Human Genome Studies, DOE Joint
 Genome Institute, Los Alamos National Laboratory, MS M888, Los
 Alamos, NM 87545, USA
 REFERENCE 2 (bases 1 to 168612)
 AUTHORS Bruce, D., Munt, M., Doggett, N., Munk, C., Saunders, E., Robinson, D.,
 Jones, D., Buckingham, J., Chasteen, L., Thompson, S., Goodwin, L.,
 Bryant, J., Tesmer, J., Meincke, L., Longmire, J., White, S., Tatum, O.,
 Campbell, C., Fawcett, J., Maltbie, M., Bussod, M., Sutherland, R.,
 McMurtry, K., Han, C. and Deaven, L.

REFERENCE
 AUTHORS Jones, D., Munt, M., Doggett, N., Munk, C., Saunders, E., Robinson, D.,
 Bryant, J., Tesmer, J., Chasteen, L., Thompson, S., Goodwin, L.,
 Campbell, C., Fawcett, J., Meincke, L., Longmire, J., White, S., Tatum, O.,
 Campbell, C., Fawcett, J., Maltbie, M., Bussod, M., Sutherland, R.,
 McMurtry, K., Han, C. and Deaven, L.
 TITLE Direct Submission
 JOURNAL Submitted (21-APR-1999) Center for Human Genome Studies, DOE Joint
 Genome Institute, Los Alamos National Laboratory, MS M888, Los
 Alamos, NM 87545, USA
 REFERENCE 3 (bases 1 to 168612)

REFERENCE
 AUTHORS Jones, D., Munt, M., Doggett, N., Munk, C., Saunders, E., Robinson, D.,
 Bryant, J., Tesmer, J., Chasteen, L., Thompson, S., Goodwin, L.,
 Campbell, C., Fawcett, J., Meincke, L., Longmire, J., White, S., Tatum, O.,
 Campbell, C., Fawcett, J., Maltbie, M., Bussod, M., Sutherland, R.,
 McMurtry, K., Han, C. and Deaven, L.
 TITLE Direct Submission
 JOURNAL Submitted (21-APR-1999) Center for Human Genome Studies, DOE Joint
 Genome Institute, Los Alamos National Laboratory, MS M888, Los
 Alamos, NM 87545, USA
 REFERENCE 4 (bases 1 to 168612)

AUTHORS

TITLE
 JOURNAL

REFERENCE
 AUTHORS

TITLE
 JOURNAL

REFERENCE
 AUTHORS

TITLE
 JOURNAL

COMMENT

FEATURES
 source
 1. 168612
 Location/Qualifiers
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="16"
 /clone="RP11-381L15"

ORIGIN

Query Match 88.9%; Score 16; DB 9; Length 168612;
 Best Local Similarity 100.0%; Pred. No. 8.6e+02;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAGAACACACAGATTC 16
 |||||
 Db 61304 GAGAACACACAGATTC 61319

RESULT 23
 AC007333 180837 bp DNA linear PRI 27-MAR-2003
 LOCUS Homo sapiens chromosome 16 clone RP11-142G1, complete sequence.
 DEFINITION
 AC007333
 ACCESSION
 VERSION AC007333.8 GI:29294008
 KEYWORDS HTG.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE
 AUTHORS Bruce, D., Munt, M., Doggett, N., Munk, C., Saunders, E., Robinson, D.,
 Jones, D., Buckingham, J., Chasteen, L., Thompson, S., Goodwin, L.,
 Bryant, J., Tesmer, J., Meincke, L., Longmire, J., White, S., Tatum, O.,
 Campbell, C., Fawcett, J., Maltbie, M., Bussod, M., Sutherland, R.,
 McMurtry, K., Han, C. and Deaven, L.
 TITLE Direct Submission
 JOURNAL Submitted (21-APR-1999) Center for Human Genome Studies, DOE Joint
 Genome Institute, Los Alamos National Laboratory, MS M888, Los
 Alamos, NM 87545, USA
 REFERENCE 2 (bases 1 to 180837)
 AUTHORS Bruce, D., Munt, M., Doggett, N., Munk, C., Saunders, E., Robinson, D.,
 Jones, D., Buckingham, J., Chasteen, L., Thompson, S., Goodwin, L.,
 Bryant, J., Tesmer, J., Meincke, L., Longmire, J., White, S., Tatum, O.,
 Campbell, C., Fawcett, J., Maltbie, M., Bussod, M., Sutherland, R.,
 McMurtry, K., Han, C. and Deaven, L.

REFERENCE
 AUTHORS Bruce, D., Munt, M., Doggett, N., Munk, C., Saunders, E., Robinson, D.,
 Jones, D., Buckingham, J., Chasteen, L., Thompson, S., Goodwin, L.,
 Bryant, J., Tesmer, J., Meincke, L., Longmire, J., White, S., Tatum, O.,
 Campbell, C., Fawcett, J., Maltbie, M., Bussod, M., Sutherland, R.,
 McMurtry, K., Han, C. and Deaven, L.
 TITLE Direct Submission
 JOURNAL Submitted (21-APR-1999) Center for Human Genome Studies, DOE Joint
 Genome Institute, Los Alamos National Laboratory, MS M888, Los
 Alamos, NM 87545, USA
 REFERENCE 3 (bases 1 to 180837)
 AUTHORS Bruce, D., Munt, M., Doggett, N., Munk, C., Saunders, E., Robinson, D.,
 Jones, D., Buckingham, J., Chasteen, L., Thompson, S., Goodwin, L.,
 Bryant, J., Tesmer, J., Meincke, L., Longmire, J., White, S., Tatum, O.,
 Campbell, C., Fawcett, J., Maltbie, M., Bussod, M., Sutherland, R.,
 McMurtry, K., Han, C. and Deaven, L.

REFERENCE
 AUTHORS Bruce, D., Munt, M., Doggett, N., Munk, C., Saunders, E., Robinson, D.,
 Jones, D., Buckingham, J., Chasteen, L., Thompson, S., Goodwin, L.,
 Bryant, J., Tesmer, J., Meincke, L., Longmire, J., White, S., Tatum, O.,
 Campbell, C., Fawcett, J., Maltbie, M., Bussod, M., Sutherland, R.,
 McMurtry, K., Han, C. and Deaven, L.
 TITLE Direct Submission
 JOURNAL Submitted (21-APR-1999) Center for Human Genome Studies, DOE Joint
 Genome Institute, Los Alamos National Laboratory, MS M888, Los
 Alamos, NM 87545, USA
 REFERENCE 4 (bases 1 to 180837)

REFERENCE
 AUTHORS Bruce, D., Munt, M., Doggett, N., Munk, C., Saunders, E., Robinson, D.,
 Jones, D., Buckingham, J., Chasteen, L., Thompson, S., Goodwin, L.,
 Bryant, J., Tesmer, J., Meincke, L., Longmire, J., White, S., Tatum, O.,
 Campbell, C., Fawcett, J., Maltbie, M., Bussod, M., Sutherland, R.,
 McMurtry, K., Han, C. and Deaven, L.
 TITLE Direct Submission
 JOURNAL Submitted (21-APR-1999) Center for Human Genome Studies, DOE Joint
 Genome Institute, Los Alamos National Laboratory, MS M888, Los
 Alamos, NM 87545, USA
 REFERENCE 5 (bases 1 to 180837)

* NOTE: This is a 'working draft' sequence. It currently consists of 5 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

* 60867: contig of 60867 bp in length
 * 60868 60967: gap of unknown length
 * 60968 20854: contig of 14887 bp in length
 * 20955 21126: gap of unknown length
 * 21127 21226: gap of unknown length
 * 21227 21275: contig of 1531 bp in length
 * 21275 21285: gap of unknown length
 * 21285 21402: contig of 1169 bp in length.

FEATURES

source

1. 214026
 /organism="Rattus norvegicus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10116"
 /clones="CH230-143J16"
 1. 1534
 /note="wgs_contig"

ORIGIN

Query Match 88.9%; Score 16; DB 2; Length 214026;
 Best Local Similarity 100.0%; Pred. No. 8.5e+02;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 GAGAACACACAGATTC 16
 Db 99497 GAGAACACACAGATTC 99512

RESULT 25
 AC094494
 LOCUS Rattus norvegicus clone CH230-4D21, *** SEQUENCING IN PROGRESS ***,
 DEFINITION 2 unordered pieces.
 AC094494
 AC094494.7 GI:304666802
 HTG: HTGS PHASE1: HTGS DRAFT: HTGS_ENRICHED.
 KEYWORDS Rattus norvegicus (Norway rat)
 SOURCE Rattus norvegicus
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.

REFERENCE 1 (bases 1 to 236172)
 AUTHORS Muzny, D., Marle, Metzker, M., Lee, Abramson, S., Adams, C., Alder, J., Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguiano, D., Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H., Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F., Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M., Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E., Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J., Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L., Davila, M.L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D., Deigado, O., Denson, S., Derano, C., Ding, Y., Dinh, H., Divya, K., Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Evans, K., Egan, A., Escotto, M., Eugene, C., Evans, C.A., Falls, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Fraser, C.M., Gabisi, A., Ganja, R., Garcia, A., Garner, T., Garza, M., Gabregeorgis, E., Geer, K., Gill, R., Grady, M., Guerra, M., Guevara, W., Gunaratne, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K., Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S.L., Hodgson, N., Hogues, M., Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A., Karpathy, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kowis, C., Kraft, C.L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J.,

Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lorensuhewa, L., Louised, H., Lorado, R.J., Lu, X., Ma, J., Maheshwari, M., Mahindaratne, M., Mahmoud, M., Malloy, K., Mangum, A., Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E., Mathew, S., McLeod, M.P., McNeill, T.Z., Meenen, E., Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwaekelemeh, O., Okunolu, G., Olariunpungoon, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfrankoch, C., Popper, F., Poindexter, A., Popovic, D., Primus, E., Pu, L.-L., Puazo, M., Quiroz, J., Rachlin, B., Reeves, K., Regier, M.A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S.J., Sanders, M., Savery, G., Scherer, S., Scott, G., Shatsman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C.D., Smajda, D., Sneed, A., Sodergren, R., Song, X.-Z., Sorelle, R., Sosa, J., Steimle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K., Valas, R., Vera, V., Villasana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wlezyk, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Niederhausern, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O., Weinstein, G. and Gibbs, R.A.

Direct Submission

Unpublished
 2 (bases 1 to 236172)

Worley, K.C.

Direct Submission

Submitted (15-SEP-2001) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

3 (bases 1 to 236172)

Rat Genome Sequencing Consortium.

Direct Submission

Submitted (09-MAY-2003) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

On May 9, 2003 this sequence version replaced gi:24942238.

The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center

Center: Baylor College of Medicine

Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>

Contact: hgsc-help@bcm.tmc.edu

----- Project Information

Center project name: GATN

Center clone name: CH230-4D21

----- Summary Statistics

Assembly program: Atlas

Consensus quality: 220342 bases at least Q40

Consensus quality: 22840 bases at least Q30

Consensus quality: 22594 bases at least Q20

Estimated insert size: 235672; sum-of-contigs estimation

Quality coverage: 6x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
 * (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
 * NOTE: This is a 'working draft' sequence. It currently consists of 2 contigs. The true order of the pieces

* is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

* 1 234509: contig of 234509 bp in length
 * 234510 234609: gap of unknown length
 * 234610 236172: contig of 1563 bp in length.

FEATURES

source

1. .236172
 /organism="Rattus norvegicus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10116"
 /clone="CH230-4D21"
 1. .1828
 /note="wgs_contig"

ORIGIN

Query Match 88.9%; Score 16; DB 2; Length 236172;

Best Local Similarity 100.0%; Pred. No. 8.5e+02;

Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAGAACACACAGATTC 16

Db 54358 GAGAACACACAGATTC 54373

RESULT 26

AC126106/c

LOCUS

DEFINITION AC126106 251032 bp DNA linear HTG 10-MAY-2003

Rattus norvegicus clone CH230-2604, *** SEQUENCING IN PROGRESS ***;

2 unordered pieces.

AC126106 GI:30522483

VERSION HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED.

KEYWORDS Rattus norvegicus (Norway rat)

SOURCE

ORGANISM

Rattus norvegicus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;

Rattus.

1 (bases 1 to 251032)

Muzny,D,Marie., Metzker,M, Lee., Abramson,S., Adams,C., Alder,J.,

Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D.,

Anylebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,

Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,

Biswal,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,

Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,

Cardenas,V., Carter,K., Cavazos,I., Caesar,H., Center,A.,

Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,

Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,

Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,

Dalgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,

Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,

Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,I., Fan,G.,

Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,

Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,

Greggeorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W.,

Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K.,

Harvey,Y., Havlak,P., Haves,A., Henderson,N., Hernandez,J.,

Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hughes,M.,

Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A.,

Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,

Karpachy,S., Kelly,S., Khan,Z., Khan,Z., King,L., Kovar,C.,

Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,

Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,

Lorenshewa,L., Loulsged,H., Lozano,R.J., Lu,X., Ma,J.,

Maheshwari,M., Mahindartine,M., Mahmoud,M., Malloy,K., Mangum,A.,

Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E.,

Mawhinney,S., McLeod,M.P., McNeill,T.Z., Meenen,E.,

Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S.,

Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L.,

Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S.,

Nwaokemele, O., Okwuonu, G., Olarnpunsagoon, A., Pal, S., Parks, K.,
 Paternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C.,
 Popper, F., Poudexter, A., Popovic, D., Primus, E., Pu, L.,
 Puzos, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M.A., Reigh, R.,
 Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F.,
 Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S.J.,
 Sanders, W., Savary, G., Scherer, S., Scott, G., Shatsman, S., Shen, H.,
 Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C.D., Smajda, D.,
 Sneed, A., Sodergren, E., Song, X., Z., Sorelle, R., Sosa, J.,
 Steimle, M., Strong, R., Sutton, A., Svatek, A., Taber, P., Taylor, C.,
 Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K.,
 Valas, R., Vera, V., Villasana, D., Waldron, L., Walker, B., Wang, J.,
 Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F.,
 Williams, G., Willson, R., Wleczyk, R., Wooden, H., Worley, K.,
 Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V.,
 Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Zhao, S., Dunn, D., von
 Niederhausern, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O.,
 Weinstock, G. and Gibbs, R.A.

Direct Submission

Unpublished

2 (bases 1 to 251032)

Worley, K.C.

Direct Submission

Submitted (03-JUL-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA

3 (bases 1 to 251032)

Rat Genome Sequencing Consortium.

Direct Submission

Submitted (10-MAY-2003) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA

On May 10, 2003 this sequence version replaced gi:23264220.

The sequence in this assembly is a combination of BAC based reads
 and whole genome shotgun sequencing reads assembled using Atlas
 (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
 in the feature table below represents a scaffold in the Atlas
 assembly (a 'contig-scaffold'). Within each contig-scaffold,
 individual sequence contigs are ordered and oriented, and separated
 by sized gaps filled with Ns to the estimated size. The sequence
 may extend beyond the ends of the clone and there may be sequence
 contigs within a contig-scaffold that consist entirely of whole
 genome shotgun sequence reads. Both end sequences and whole genome
 shotgun sequence only contigs will be indicated in the feature
 table.

Genome Center

Center: Baylor College of Medicine

Center code: BCM

Web site: http://www.hgsc.bcm.tmc.edu/

Contact: hgsc-help@bcm.tmc.edu

----- Project Information

Center project name: GQCN

Center clone name: CH230-2604

----- Summary Statistics

Assembly program: Atlas 3.0;

Consensus quality: 243807 bases at least Q40

Consensus quality: 249909 bases at least Q30

Consensus quality: 247143 bases at least Q20

Estimated insert size: 259832; sum-of-contigs estimation

Quality coverage: 6x in Q20 bases; sum-of-contigs estimation

 * NOTE: Estimated insert size may differ from sequence length
 * (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html)
 * NOTE: This sequence may represent more than one 'clone'.
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 2 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

* 1 249354: contig of 249354 bp in length

* 249355 249454: gap of unknown length
 * 249455 251032: contig of 1578 bp in length.

FEATURES

source
 1. 251032
 /organism="Rattus norvegicus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10116"
 /clone="CH230-2604"
 743. .1295
 /note="clone_boundary
 clone end:T7
 site:ECORI
 end_sequence:BH334758"
 complement(245981..246654)
 /note="clone_boundary
 clone end:Sp6
 site:ECORI
 end_sequence:BH334759"

misc_feature

misc_feature

ORIGIN

Query Match 88.9%; Score 16; DB 2; Length 251032;

Best Local Similarity 100.0%; Pred. No. 8.5e+02;

Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GAGAACCAACAGATTC 16

Db 163240 GAGACACACAGATTC 163225

RESULT 27

AC094225

LOCUS Rattus norvegicus clone CH230-3B10, WORKING DRAFT SEQUENCE, 3
 DEFINITION unordered pieces.

ACCESSION

AC094225.6 GI:30467636

HTG; HTGS PHASE1; HTGS DRAFT; HTGS_FULLTOP.

KEYWORDS Rattus norvegicus (Norway rat)

SOURCE

ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.

1 (bases 1 to 292516)

REFERENCE

AUTHORS

Muzny, D., Maric, M., Metzker, M., Lee, S., Abramson, S., Adams, C., Alder, J.,
 Allen, C., Allen, H., Albrooks, S., Amin, A., Anguiano, D.,
 Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H.,
 Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F.,
 Biswal, K., Blair, J., Blankenburg, K., Blych, P., Brown, M.,
 Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E.,
 Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A.,
 Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J.,
 Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L.,
 Davila, M., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D.,
 Delgado, O., Denson, S., Derano, C., Ding, Y., Dinh, H., Divya, K.,
 Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K.,
 Egan, A., Escotto, M., Eugene, C., Evans, C. A., Falls, T., Fan, G.,
 Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P.,
 Fraser, C. M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M.,
 Gregeorge, E., Geer, K., Gill, R., Grady, M., Guerra, M., Guevara, W.,
 Gunaratne, P., Haaland, W., Hamil, C., Hamilton, C., Hamilton, K.,
 Harvey, Y., Havlak, P., Hayes, A., Henderson, N., Hernandez, J.,
 Hernandez, R., Hines, S., Hladun, S. L., Hodgson, A., Hogue, M.,
 Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A.,
 Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A.,
 Karpethy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C.,
 Kowis, C., Kraft, C. L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J.,
 Liu, J., Liu, W., Liu, W., London, P., Longacre, S., Lopez, J.,
 Lorensuhewa, L., Louleghed, H., Lozado, R. J., Lu, X., Ma, J.,
 Maheshwari, M., Mahindartne, M., Mahmoud, M., Malloy, K., Mangum, A.,
 Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E.,
 Mathew, S., McLeod, M. P., McNeill, T. Z., Meenen, E.,
 Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S.,
 Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L.,

Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S.,
 Nwaokeme, O., Okwunnu, G., Olarnpunsagoon, A., Pal, S., Parks, K.,
 Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C.,
 Plopper, F., Poindexter, A., Popovic, D., Primus, E., Pu, L.,
 Puaro, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M. A., Reigh, R.,
 Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F.,
 Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S. J.,
 Sanders, W., Savary, G., Scherer, S., Scott, G., Shatsman, S., Shen, H.,
 Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C. D., Snajls, D.,
 Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J.,
 Steimle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C.,
 Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K.,
 Valas, R., Vera, V., Villanueva, D., Waldron, L., Walker, B., Wang, J.,
 Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F.,
 Williams, G., Willson, R., Wlezyk, R., Woodson, H., Wortley, K.,
 Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V.,
 Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von
 Niederhausern, A., Weiss, R., Smith, D. R., Holt, R. A., Smith, H. O.,
 Weinstein, G. and Gibbs, R. A.

Direct Submission

Unpublished

2 (bases 1 to 292516)

Worley, K. C.

Direct Submission

JOURNAL

TITLE

AUTHORS

JOURNAL

REFERENCE

AUTHORS

JOURNAL

TITLE

JOURNAL

COMMENT

JOURNAL

TITLE

AUTHORS

JOURNAL

TITLE

AUTHORS

JOURNAL

TITLE

AUTHORS

JOURNAL

TITLE

AUTHORS

JOURNAL

TITLE

AUTHORS

JOURNAL

TITLE

AUTHORS

JOURNAL

TITLE

AUTHORS

JOURNAL

TITLE

AUTHORS

JOURNAL

TITLE

AUTHORS

JOURNAL

TITLE

AUTHORS

JOURNAL

TITLE

AUTHORS

JOURNAL

TITLE

AUTHORS

JOURNAL

TITLE

AUTHORS

JOURNAL

TITLE

AUTHORS

JOURNAL

```

* 1 287193: contig of 287193 bp in length
* 287194 287293: gap of unknown length
* 287294 288662: contig of 1369 bp in length
* 288663 288762: gap of unknown length
* 288763 292516: contig of 3754 bp in length.
FEATURES
    source
        1..292516
            /organism="Rattus norvegicus"
            /mol_type="genomic DNA"
            /db_xref="taxon:10116"
            /clone="CH230-3B10"
        1..2212
            /note="wgs contig"
        45578..48328
            /note="wgs contig"
        complement(285431..286059)
            /note="clone boundary"
            clone end: r7
            site: EcoRI
            end_sequence: BH307153"

ORIGIN
Query Match      88.9%; Score 16; DB 2; Length 292516;
Best Local Similarity 100.0%; Pred. No. 8.5e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAGACACACACAGATTC 16
    |||||
Db 135976 GAGACACACAGATTC 135991

RESULT 28
AX340908
LOCUS AX340908 485 bp DNA linear PAT 10-JAN-2002
DEFINITION Sequence 1155 from Patent WO0196388.
ACCESSION AX340908
VERSION AX340908.1 GI:18136890
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Jiang, Y., Harlocker, S.L. and Secrist, H.
TITLE Compositions and methods for the therapy and diagnosis of colon
cancer
JOURNAL Patent: WO 0196388-A 1155 20-DEC-2001;
CORIXA CORPORATION (US)
FEATURES
    source
        1..485
            /organism="Homo sapiens"
            /mol_type="unassigned DNA"
            /db_xref="taxon:9606"

ORIGIN
Query Match      85.6%; Score 15.4; DB 6; Length 485;
Best Local Similarity 94.1%; Pred. No. 2.2e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 AGAACACACAGATTCGC 18
    |||||
Db 48 AGCACACACAGATTCGC 64

RESULT 29
AR396508
LOCUS AR396508 534 bp DNA linear PAT 18-DEC-2003
DEFINITION Sequence 2523 from patent US 6617156.
ACCESSION AR396508
VERSION AR396508.1 GI:40126069
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.

Unclassified.
REFERENCE 1 (bases 1 to 534)
AUTHORS Doucette-Stamm, L.A. and Bush, D.
TITLE Nucleic acid and amino acid sequences relating to Enterococcus faecalis for diagnostics and therapeutics
JOURNAL Patent: US 6617156-A 2523 09-SEP-2003;
FEATURES
    Location/Qualifiers
        source
            1..534
                /organism="unknown"
                /mol_type="genomic DNA"

ORIGIN
Query Match      85.6%; Score 15.4; DB 6; Length 534;
Best Local Similarity 94.1%; Pred. No. 2.2e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 AGAACACACAGATTCGC 18
    |||||
Db 465 AGAACACACAGATTCGC 481

RESULT 30
AY344560/c
LOCUS AY344560 886 bp DNA linear VRT 01-AUG-2004
DEFINITION Sinogastromyzon puliensis haplotype tw4 control region, complete
sequence; mitochondrial.
ACCESSION AY344560
VERSION AY344560.1 GI:38098435
KEYWORDS
SOURCE mitochondrion Sinogastromyzon puliensis
ORGANISM Sinogastromyzon puliensis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
Cypriniformes; Balitoridae; Balitorinae; Sinogastromyzon.
REFERENCE 1 (bases 1 to 886)
AUTHORS Liao, T.-Y., Wang, T.-Y., Shen, S.-J. and Tzeng, C.-S.
TITLE The Effects of Glaciation on the Dispersion Pattern of the
Endangered Species, Sinogastromyzon puliensis (Cypriniformes:
Balitoridae), in Southwest Taiwan
JOURNAL Unpublished
AUTHORS Liao, T.-Y., Wang, T.-Y., Shen, S.-J. and Tzeng, C.-S.
TITLE Direct Submission
JOURNAL Submitted (17-JUL-2003) Life Science, NTHU, 101, Section 2 Kuang Fu
Road, Hsinchu, Taiwan 300, Republic of China
FEATURES
    source
        1..886
            /organism="Sinogastromyzon puliensis"
            /organelle="mitochondrion"
            /mol_type="genomic DNA"
            /db_xref="taxon:241472"
            /haplotype="tw4"
        1..886
            /note="control region"

ORIGIN
Query Match      85.6%; Score 15.4; DB 5; Length 886;
Best Local Similarity 94.1%; Pred. No. 2.2e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GAGAACACACAGATTCGC 17
    |||||
Db 177 GAGAACACACATATTCG 161

RESULT 31
AY344557/c
LOCUS AY344557 887 bp DNA linear VRT 01-AUG-2004
DEFINITION Sinogastromyzon puliensis haplotype tw1 control region, complete
sequence; mitochondrial.
ACCESSION AY344557
VERSION AY344557.1 GI:38098432
KEYWORDS

```



```

JOURNAL      Submitted (17-JUL-2003) Life Science, NTHU, 101, Section 2 Kuang Fu
FEATURES
source       Road, Hsinchu, Taiwan 300, Republic of China
              Location/Qualifiers
                1..887
                /organism="Sinogastromyzon puliensis"
                /organelle="mitochondrion"
                /mol_type="genomic DNA"
                /db_xref="taxon:241472"
                /haplotype="tw6"
              misc_feature 1..887
                /note="control region"
ORIGIN
Query Match      85.6%;   Score 15.4;   DB 5;   Length 887;
Best Local Similarity 94.1%;   Pred. No. 2.2e+03;
Matches 16;   Conservative 0;   Mismatches 1;   Indels 0;   Gaps 0;

QY      1  GAGAACAAACAAGATTTCG 17
        |||||
        178 GAGAACAAACAATATTCG 162

RESULT 35
AY344558/c
LOCUS      AY344558
DEFINITION Sinogastromyzon puliensis haplotype tw2 control region, complete
            sequence; mitochondrial.
ACCESSION  AY344558
VERSION    AY344558.1 GI:38098433
KEYWORDS   .
SOURCE     Mitochondrion Sinogastromyzon puliensis
ORGANISM   Sinogastromyzon puliensis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
            Cypriniformes; Balitoridae; Balitorinae; Sinogastromyzon.
            1. (bases 1 to 888)
REFERENCE  Liao,T.-Y., Wang,T.-Y., Shen,S.-J. and Tzeng,C.-S.
            The Effects of Glaciation on the Dispersion Pattern of the
            Endangered Species, Sinogastromyzon puliensis (Cypriniformes:
            Balitoridae), in Southwest Taiwan
            Unpublished
REFERENCE  2.(bases 1 to 888)
            Liao,T.-Y., Wang,T.-Y., Shen,S.-J. and Tzeng,C.-S.
            Direct Submission
            Submitted (17-JUL-2003) Life Science, NTHU, 101, Section 2 Kuang Fu
            Road, Hsinchu, Taiwan 300, Republic of China
FEATURES
source       1..888
              /organism="Sinogastromyzon puliensis"
              /organelle="mitochondrion"
              /mol_type="genomic DNA"
              /db_xref="taxon:241472"
              /haplotype="tw2"
            misc_feature 1..888
              /note="control region"
ORIGIN
Query Match      85.6%;   Score 15.4;   DB 5;   Length 888;
Best Local Similarity 94.1%;   Pred. No. 2.2e+03;
Matches 16;   Conservative 0;   Mismatches 1;   Indels 0;   Gaps 0;

QY      1  GAGAACAAACAAGATTTCG 17
        |||||
        178 GAGAACAAACAATATTCG 162

RESULT 36
JSRTMPB
LOCUS      JSRTMPB
DEFINITION Jaggsakte sheep retrovirus pol protein (pol) gene, partial cds.
ACCESSION  M65014
VERSION    M65014.1 GI:331345
KEYWORDS   pol polyprotein.

```

Kusumegi, T., Oka, M., Ryu, R., Ueda, M., Matsubara, K., RIKEN;
 Kawai, J., Carninci, P., Adachi, J., Aizawa, K., Arakawa, T., Fukuda, S.,
 Hara, A., Hashidume, W., Hayatsu, N., Imorani, K., Ishii, Y., Itoh, M.,
 Kagawa, I., Kondo, S., Konno, H., Miyazaki, A., Osato, N., Ota, Y.,
 Saito, R., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T.,
 Yoshino, M., and Hayashizaki, Y.
 Collection, mapping, and annotation of over 28,000 cDNA clones from
 japonica rice

Science 301 (5631), 376-379 (2003)
 22752273
 12869764

2 (bases 1 to 1664)

Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Doi, K.,
 Fujimura, T., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W.,
 Hayashida, K., Hayashizaki, Y., Hayatsu, N., Hiramoto, K., Hiraoka, T.,
 Hori, F., Hotta, I., Iida, Y., Ikeda, R., Imamura, K., Itoh, M.,
 Inotani, K., Ishibiki, J., Ishii, Y., Ishikawa, M., Itoh, M., Kagawa, I.,
 Kanagawa, S., Katoh, H., Kawagashira, N., Kawai, J., Kawamata, M.,
 Kikuchi, S., Kishikawa-Hirozane, T., Kishimoto, N., Kobayashi, M.,
 Kodama, T., Kojima, K., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,
 Koya, S., Kurihara, C., Kurosaki, T., Kusumegi, T., Li, C., Lu, M.,
 Masuda, H., Matsubara, K., Matsuyama, T., Miura, J., Miyazaki, A.,
 Mizuno, K., Murakami, K., Murata, M., Nagata, T., Nakamura, M.,
 Namiki, T., Narikawa, R., Niikura, J., Nishi, K., Nomura, K.,
 Numasaki, R., Oneda, E., Ohno, M., Ohtsuki, K., Oka, M., Ooka, H.,
 Osato, N., Ota, Y., Otomo, Y., Ryu, R., Saitoh, H., Sakai, C., Sakai, K.,
 Sakazume, N., Sano, H., Sasaki, D., Sato, K., Satoh, K., Shibata, K.,
 Shinagawa, A., Shiraki, T., Shishiki, T., Sogabe, Y., Sugano, S.,
 Sugiyama, A., Suzuki, K., Suzuki, Y., Tagami, M., Tagami-Takeda, Y.,
 Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A.,
 Toya, T., Tsunoda, Y., Ueda, M., Waki, K., Xie, Q., Yahagi, W.,
 Yanada, H., Yamamoto, M., Yasunishi, A., Yazaki, J., Yokomizo, S. and
 Yoshimura, A.

Direct Submission
 Submitted (27-AUG-2002) Shoshi Kikuchi, National Institute of
 Agrobiological Sciences, Department of Molecular Genetics, Head of
 Laboratory of Gene Expression; 2-1-2 Kanonndai, Tsukuba, Ibaraki
 305-8602 Japan (E-mail: skikuchi@nias.affrc.go.jp,
 Tel: 81-29-838-7007, Fax: 81-29-838-7007)

This clone is one of the 28K full-length cDNA clones from japonica
 rice.

URL : <http://cdna01.dna.affrc.go.jp/cDNA/>
 NIAS Rice Full-length cDNA Project Team: Kikuchi, S., Satoh, K.,
 Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J.,
 Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Namiki, T.,
 Oneda, E., Yahagi, W., Suzuki, K., Li, C., Ohtsuki, K., Shishiki, T. and
 Yamamoto, M.

FAIS Genome Sequencing & Analysis Group: Otomo, Y., Iida, Y.,
 Fujimura, T., Ikeda, R., Ishibiki, J., Kawamata, M., Kobayashi, M.,
 Kodama, T., Kurosaki, T., Kusumegi, T., Lu, M., Masuda, H., Miura, J.,
 Mizuno, K., Narikawa, R., Niikura, J., Oka, M., Ryu, R., Sugano, S.,
 Sugiyama, A., Suzuki, Y., Tsunoda, Y., Ueda, M., Xie, Q., Yokomizo, S.,
 Yoshimura, A., Matsubara, K. and Murakami, K.
 Genome Exploration Research Group in Riken Genomic Sciences Center
 and Genome Science Laboratory in Riken: Adachi, J., Aizawa, K.,
 Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hanagaki, T.,
 Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K.,
 Hiraoka, T., Hori, F., Iida, J., Imamura, K., Inotani, K., Ishii, Y.,
 Itoh, M., Kagawa, I., Kanagawa, S., Katoh, H., Kawai, J., Kouda, M.,
 Kishikawa-Hirozane, T., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,
 Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,
 Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Ohno, M., Osato, N.,
 Ota, Y., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H.,
 Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T.,
 Sogabe, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F.,
 Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T., Waki, K.,
 Yasunishi, A. and Hayashizaki, Y.

Location/Qualifiers

1. 1664
 /organism="Oryza sativa (japonica cultivar-group)"
 /mol_type="mRNA"
 /cultivar="Nipponbare"
 /db_xref="taxon:39947"

ORIGIN
 Query Match 85.6%; Score 15.4; DB 8; Length 1664;
 Best Local Similarity 94.1%; Pred. No. 2.2e+03;
 Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 AGAACCAACAAAGATTGCG 18

||||| ||||||| |||||

Db 987 AGRACCTACAGATTGCG 971

RESULT 38

AY692867

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

mrna

CDS

FEATURES

source

mrna

CDS

FEATURES

source

mrna

CDS

FEATURES

source

mrna

CDS

FEATURES

source

mrna

CDS

FEATURES

source

mrna

CDS

FEATURES

source

mrna

CDS

FEATURES

source

mrna

CDS

FEATURES

source

mrna

CDS

FEATURES

source

mrna

CDS

FEATURES

source

mrna

CDS

FEATURES

source

mrna

CDS

FEATURES

/clone="002-145-B01"
 Query Match 85.6%; Score 15.4; DB 8; Length 1664;
 Best Local Similarity 94.1%; Pred. No. 2.2e+03;
 Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 2 AGAACCAACAAAGATTGCG 18
 ||||| ||||||| |||||
 Db 987 AGRACCTACAGATTGCG 971
 AY692867 1950 bp DNA linear PLN 11-AUG-2004
 Saccharomyces cerevisiae clone FLH158310.01X YIL155C gene, complete
 cds
 AY692867 GI:51013184
 Yeast ORF Project.
 Saccharomyces cerevisiae (baker's yeast)
 Saccharomyces cerevisiae
 Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 1 (bases 1 to 1950)
 Marsischky, G., Rolfs, A., Richardson, A., Kane, M., Baqui, M.,
 Taycher, E., Hu, Y., Vannberg, F., Weger, J., Kramer, J., Moreira, D.,
 Kelley, F., Zuo, D., Raphael, J., Hogle, C., Jepson, D., Williamson, J.,
 Canarso, A., Gonzaga, L., Vasconcelos, A.T., Simpson, A., Kolodner, R.,
 Harlow, E. and Labaer, J.
 Creation of the YFLEX clone resource: cloning of Saccharomyces
 cerevisiae ORFs in the Gateway recombinational cloning system
 Unpublished
 2 (bases 1 to 1950)
 Marsischky, G., Rolfs, A., Richardson, A., Kane, M., Baqui, M.,
 Taycher, E., Hu, Y., Vannberg, F., Weger, J., Kramer, J., Moreira, D.,
 Kelley, F., Zuo, D., Raphael, J., Hogle, C., Jepson, D., Williamson, J.,
 Canarso, A., Gonzaga, L., Vasconcelos, A.T., Simpson, A., Kolodner, R.,
 Harlow, E. and Labaer, J.
 Direct Submission
 Submitted (20-JUL-2004) Biological Chemistry and Molecular
 Pharmacology, Harvard Institute of Proteomics, 320 Charles St.,
 Cambridge, MA 02141, USA
 This clone is part of a collection of Saccharomyces cerevisiae
 full-length ORF clones generated by the Harvard Institute of
 Proteomics. Each CDS has been cloned with its native stop codon.
 The CDS has been directionally cloned using the Gateway cloning
 system into the donor vectors pDONR 201 or pDONR 221. Additional
 sequences in the clone: 'TCCAGCTGACCAC' after the attL1 site and
 before the 'ATG' (from Research Genetics primers used to amplify
 the ORFs, including a Kozak consensus sequence).
 'ATCCCGGGAATTCCTATG' after the stop codon and before the attL2
 site (from the Research Genetics primers used to amplify the ORFs).
 Location/Qualifiers
 1. 1950
 /organism="Saccharomyces cerevisiae"
 /mol_type="genomic DNA"
 /db_xref="taxon:4932"
 /clone="FLH158310.01X"
 /lab_host="DH5alpha T1 resistant"
 <1. 1950
 /product="YIL155C"
 1. 1950
 /codon_start=1
 /product="YIL155C"
 /protein_id="AAT92886.1"
 /db_xref="GI:51013185"

source

mrna

CDS

FEATURES

source

mrna

CDS

FEATURES

source

mrna

CDS

FEATURES

source

mrna

CDS

FEATURES

source

mrna

CDS

FEATURES

source

mrna

CDS

FEATURES

translation="NFSVTRRAAGAAAAMATATGTYWVTSGDRPLVNDPSYVQV
 FPTAAPQVSRDLRLKATQHFVLIIGGATGTCALDAATGALNVALVEKGDFA
 SGTSSKTMHGGVRYLEKAFWEFKQAQDLVIEALNERKILINTAPHLCTVPLLI
 PYSTWQVPIYMGCKFPYDFAGSOKLSKSYLKSATVEKAPMLTNDLKASLVVHD
 GSFNDRNLNATLAI TAVENGATVLYNVEVORLIKDPTSGKVI GAARDVETNELVRIN
 AKCVNATGPGYSDAIIQMDRNPGLPDSPLNDNSKIKSTFNQIAVMDPKWVIPSIGVH

IVLPSFYCPKDMGLLDVRTSDGRVMFFLPWQGVKVLGTTDIPKQVNPENMPTEADIQ
DILKELQHYIFPVKREDVLSAWGVPLVRDPTIPADGKKGSGATQGVVRSHPFTS
DNGLITAGGKWTYRQMAESTVDKXVGVGFHNLKPCHTEDIKLAGEEWTQNYVAL
LAQNYHLGSKSNLIVQNGYRSGSIIIEFFKESMENKLPISLADKENNVIYSBENNL
VNEFTFRYFTTIGELKYSMQYECYKTPDLDFILRTRFAFLDDAKALNAVAHATVKWMD
746
/note="compared to SGD sequence"
/replace="a"

variation

ORIGIN

Query Match 85.6%; Score 15.4; DB 8; Length 1950;
Best Local Similarity 94.1%; Pred. No. 2.2e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 AGAACAAAGATTCCG 18
|||||
Db 1795 AGAACAAAGATTCCG 1811
|||||

RESULT 39
AB091267
LOCUS AB091267 1997 bp mRNA linear PLN 04-MAR-2003
DEFINITION Marchantia polymorpha mRNA for ftsZ1, complete cds.
ACCESSION AB091267
VERSION AB091267.1 GI:28804589
KEYWORDS Marchantia polymorpha (liverwort)
ORGANISM Marchantia polymorpha
Eukaryota; Viridiplantae; Streptophyta; Embryophyta;
Marchantiophyta; Marchantiopsida; Marchantiidae; Marchantiales;
Marchantiaceae; Marchantiaceae; Marchantia.

REFERENCE
Araki, Y., Takio, S., Ono, K. and Takano, H.
Two types of plastid ftsZ genes in liverwort Marchantia polymorpha
Unpublished
2 (bases 1 to 1997)
Direct Submission
Takano, H., Araki, Y., Takio, S. and Ono, K.
Submitted (03-SEP-2002) Hirooyoshi Takano, kumamoto university,
Faculty of Science, Kurokami, kumamoto-city, kumamoto 860-8555,
Japan (E-mail: takano@kumamoto-u.ac.jp).
URL: <http://www.sci.kumamoto-u.ac.jp/bio.iden/takano/english.html>,
Tel: 81-96-342-3432, Fax: 81-96-342-3432)

FEATURES
source
1..1997
/organism="Marchantia polymorpha"
/mol_type="mRNA"
/db_xref="taxon:3197"
1..1997
/gene="ftsZ1"
234..1574
/gene="ftsZ1"
/codon_start=1
/product="ftsZ1"
/protein_id="BAC57993.1"
/db_xref="GI:28804590"
/translation="MTVQVRPLTNSCHGSPVFASARPAICSGVGFARCCALPS
LHGKVSLSPLFSGTQHQDKRAEWQAKRSAGRTIRFASMI PMDSARIKXVIGVG
GGNNAINRMI GSGVGFQVFAINTDAQALLQSAATHRVQIGETLIRGLGTGNPELG
EXAAEBSLEIAEAVSDADLVFTAGMGGGTGAAPVVARLAKEGGQITGVGVTFP
TFEGRRAOQGLEAIBOLKKNVDTLIVIPNDELDDVYQATPLQEAFLIADLVLRQGV
QGISDITITPGLVNVDADVKAVMNSGTMLGVMSGTGKRAEEAQAQATSAPLIER
SIERATGVYNTITGKDLTLQEVNRVSQVVTGLADPAANIIFGAVVDEKYGAVHTII
IATGFSQTQFKTLIDPKVARQQQSPKGVDSFWRKRPAPVSRFPQGLSGKGLF"

ORIGIN

Query Match 85.6%; Score 15.4; DB 8; Length 1997;
Best Local Similarity 94.1%; Pred. No. 2.1e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 AGAACAAAGATTCCG 18
|||||

Db 1472 AGAACAAAGATTCCG 1488
|||||

RESULT 40
BD095617/c
LOCUS BD095617 2186 bp DNA linear PAT 27-AUG-2002
DEFINITION Process for biologically producing L-pipecolic acid.
ACCESSION BD095617
VERSION BD095617.1 GI:22641205
KEYWORDS WO 0148216-A/8.
SOURCE unidentified
ORGANISM unidentified
unclassified.

REFERENCE
1 (bases 1 to 2186)
AUTHORS Fujii, T., Aritoku, Y., Mukaiharu, M., Narita, T., Agematsu, H. and
Isshiki, K.
TITLE Process for biologically producing L-pipecolic acid
JOURNAL Patent: WO 0148216-A 8 05-JUL-2001;
MERCIAN CORP, TADASHI FUJII, YASUHIRO ARIKOKU, MANABU MUKAIHARA, TAKAO NARITA,
NARITA, HITOSHI AGEMATSU, KUNIO ISSHIKI
OS Escherichia coli
PN WO 0148216-A/8
PD 05-JUL-2001
PF 22-DEC-2000 WO 2000JP009137
PR 28-DEC-1999 JP 99P 373389
PI TADASHI FUJII, YASUHIRO ARIKOKU, MANABU MUKAIHARA, TAKAO NARITA,
PI HITOSHI AGEMATSU, KUNIO ISSHIKI
PC C12N15/54, C12N1/21, C12P17/12, C12N9/10, C12N15/54, C12R1/20,
PC C12N1/21, C12R1/19, C12N1/21, C12R1/15, C12P17/12, C12R1/19,
PC C12P17/12, C12R1/15, C12N9/10, C12R1/20
CC Process for biologically producing L-pipecolic acid FH Key
Location/Qualifiers
FT source 1..2186
FT /organism="Escherichia coli".
Location/Qualifiers
1..2186
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"

ORIGIN

Query Match 85.6%; Score 15.4; DB 6; Length 2186;
Best Local Similarity 94.1%; Pred. No. 2.1e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 AGAACAAAGATTCCG 18
|||||

Db 1281 AGAACAAAGATTCCG 1265
|||||

RESULT 41
CQ716675
LOCUS CQ716675 2845 bp DNA linear PAT 03-FEB-2004
DEFINITION Sequence 2609 from Patent WO02068579.
ACCESSION CQ716675
VERSION CQ716675.1 GI:42277532
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
1
AUTHORS Venter, C.J., Adams, M.C., Li, P.W. and Myers, E.W.
TITLE Kits, such as nucleic acid arrays, comprising a majority of
human exons or transcripts, for detecting expression and other uses
thereof
Patent: WO 02068579-A 2609 06-SEP-2002;
PE Corporation (NY) (US)
Location/Qualifiers
1..2845
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Query Match 85.6%; Score 15.4; DB 8; Length 1997;
Best Local Similarity 94.1%; Pred. No. 2.1e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 AGAACAAAGATTCCG 18
|||||

FT source 1. .3178 /organism='Unidentified'.
 FT Location/Qualifiers
 1. .3178
 /organism='unidentified'
 /mol_type='genomic DNA'
 /db_xref='taxon:32644'

ORIGIN
 Query Match 85.6%; Score 15.4; DB 6; Length 3178;
 Best Local Similarity 94.1%; Pred. No. 2.1e+03;
 Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 2 AGAACACACAGATTCGC 18
 |||||
 |||||
 Db 2154 AGAAGAACACAGATTCGC 2170

RESULT 46
 SCGUT2A
 LOCUS SCGUT2A 3178 bp DNA linear PLN 24-NOV-1993
 DEFINITION S.cerevisiae GUT2 gene.
 ACCESSION X71660
 VERSION X71660.1 GI:297115
 KEYWORDS glycerol-3-phosphate dehydrogenase.
 SOURCE Saccharomyces cerevisiae (baker's yeast)
 ORGANISM Saccharomyces cerevisiae
 Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 Saccharomycetales; Saccharomycetaceae; Saccharomyces.

REFERENCE
 1 Ronnow, B. and Kiehlbrandt, M. C.
 GUT2, a gene for mitochondrial glycerol 3-phosphate dehydrogenase
 of Saccharomyces cerevisiae
 JOURNAL Yeast 9 (10), 1121-1130 (1993)
 MEDLINE 94078674
 PUBMED 8256521

REFERENCE
 2 (bases 1 to 3178)
 Roennow, B.
 Direct Submission
 Submitted (26-APR-1993) B. Roennow, Carlsberg Laboratory, Dept. of
 Physiology, Gamle Carlsberg Vej 10., DK-2300 Copenhagen Valby,
 DENMARK

FEATURES
 source
 1. .3178
 /organism='Saccharomyces cerevisiae'
 /mol_type='genomic DNA'
 /strain='GRF88'
 /db_xref='taxon:4932'
 356..360
 468..2312
 /gene='GUT2'
 468..2312
 /gene='GUT2'
 /EC number='1.1.99.5'
 /codon_start=1
 /product='glycerol-3-phosphate dehydrogenase'
 /protein_id='CAA50652.1'
 /db_xref='GI:297116'
 /db_xref='GOA:P32191'
 /db_xref='Swiss-Prot:P32191'
 /translation='MTRATWCNSPPPHROVRSRDLRLDKTHQFDVLIIGGATGCT
 GCALDAATRGALNVALVEKGPASGTSKSTKMHHGGVRYLEKAFWFEKSKALDLVLEA
 LNERKHLMTATPHLCTVLPILPIYTSQWPDYIMGCKFYDFPGGSONLKKSYLLSDP
 ATVEKPMPLTDLNLKASLVHHDGFSNDRLNATLITGVGKQFVLIYEVQKLLKDP
 TSGVIGAEARDVETNLRIKACVNAVATGYSDAILOMDRNPGLSPGLDNLNSKI
 KSTFNQISLQMPKMWIPSIQVHIVLPFSYFKMGLLDVTRSDGVNMFPLPWQGVKLA
 GTTDILPKQVPEVNPMPTEADIQDILKELQHVIEFPKVRDEVLSAWAGVRLVDRDPTI
 PADDKKGATQGVVRSHEFLFTSDNGLITTAGKWTTVYRMAEETVDKVVGVGGFPHNLK
 PCHTRDYLKAGAEWTONYVALLAQNHLSKSNMNYQNYGTFESSII CEFFKESMEN
 KLPFJDAKEALNNVLYSSENNLVNFDTRFYPTFTELGKYNQVHCYRTPLDLFLLRIR
 FAFJDAKEALNNVATVTKVMGDEFNFWSEKKRQWLEKTVNFIQGRFGV'
 468..527
 /gene='GUT2'
 transit_peptide

CDS

1517. .2986
/gene="lysp"
/functions="lysine transport"
/note="resistant to lysine analogue thiosine"
/codon_start=1
/transl_table=11
/product="lysine specific permease"
/protein_id="AA17053.1"
/db_xref="GI:466778"
/translation="MSEYKTEAPGLRELKARHTMTAIGSGICTGLFVSGATIS
QAGPGALSYMLGIMVFLMTSLGELAAVPSGFATYQNYVEGFGFALGWN
WYNAVTIADVLVAGLWVSWFDPDTPGIMWSALFLGVIFLLNYSVRGFGAEYWF
LKVTVTVPIIVGLMIIGIPKAQAPAGSNWTIGEPFAGGFAAMIGAMIVGFSF
QGTGLIAGAESEDPKAPNIPRAVQVFWRIILFYFAILLISLIPIYDPSLRNDV
KDISVPTAVFOHAGLISAAVAVILTAVLISAGNSGMVASTRLMTLACDCKAPR
IFAKLSGGVPRNALTATVIAGLCFLTSWFGNQTLYLWLLNTSGMTGFIAWLGIAIS
HYFRGVIQGDHINDIPYRSGFFPGIPFIFILCLITLQNGVEAFUKDTIDWGGV
AATYIGLPLLIIFGKLIKGTFFVYSEMFPQNDKX"
3028. .3043
/gene="lysp"
/note="rho-independent"
3280. .>3655
/gene="cfr"
3280. .>3655
/gene="cfr"
/codon_start=1
/transl_table=11
/product="cfr"
/protein_id="AA17054.1"
/db_xref="GI:466779"
/translation="MRLAPFVRVGLCLSAISCAHPVLAVDDGETMVTASVQNL
KDPASISVITQDLQKVPQNLKYLKEVPQVLTNEGDKGVSIRGLSDXYLLI
VDGRVNSRNAVREHDFDLNWI"

ORIGIN

Query Match 85.6%; Score 15.4; DB 1; Length 3655;
Best Local Similarity 94.1%; Pred. No. 2.1e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 AGAACACAGATTCCG 18

|||||

Db 2273 AGAACACAGATTCCG 2257

RESULT 48

AC017735/c
LOCUS
DEFINITION
AC017735
AC017735
VERSION
AC017735.1 GI:6554264
HTG; HTGS_PHASE2
Drosophila melanogaster (fruit fly)
Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 3831)
Adams, M. and Venter, J. C.
Direct Submission
Submitted (10-DEC-1999) Celera Genomics, 45 West Gude Drive,
Rockville, MD, USA
This sequence was identified as CD1:10211919 by the submitter.
For more information on this record e-mail to fly@celera.com.
* NOTE: This is a 'working draft' sequence.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
Location/Qualifiers
1. .3831
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"

FEATURES

source

ORIGIN

Query Match 85.6%; Score 15.4; DB 2; Length 3831;
Best Local Similarity 94.1%; Pred. No. 2.1e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GAGAACACAGATTCCG 17

|||||

Db 3359 GAGAACACAGATTCCG 3343

RESULT 49

AC0911102

LOCUS

AB0911102

DEFINITION

Marchantia polymorpha gene for ftsz1, complete cds.

AC0911102

AC0911102.1 GI:28804575

KEYWORDS

SOURCE

Marchantia polymorpha (liverwort)

ORGANISM

Marchantia polymorpha

Eukaryota; Viridiplantae; Streptophyta; Embryophyta;
Marchantiophyta; Marchantiopsida; Marchantiales;
Marchantiaceae; Marchantiaceae; Marchantia.

REFERENCE

1 Araki, Y., Takio, S., Ono, K. and Takano, H.
Two types of plastid ftsz genes in liverwort Marchantia polymorpha
Unpublished
2 (bases 1 to 3926)
Takano, H., Araki, Y., Takio, S. and Ono, K.
Direct Submission
Submitted (30-AUG-2002) Hiroyoshi Takano, Kumamoto university,
Faculty of Science, Kurokami, Kumamoto-city, Kumamoto 860-8555,
Japan (E-mail: takano@kumamoto-u.ac.jp,
URL: http://www.sci.kumamoto-u.ac.jp/bio.iden/takano/english.html,
Tel: 81-96-342-3432, Fax: 81-96-342-3432)

FEATURES

Location/Qualifiers
1. .3926
/organism="Marchantia polymorpha"
/mol_type="genomic DNA"
/db_xref="taxon:3197"
join(93. .446,1536. .1635,1957. .2186,2422. .2589,2850. .3083,
3335. .3589)
/gene="fts21"
join(93. .446,1536. .1635,1957. .2186,2422. .2589,2850. .3083,
3335. .3589)
/gene="fts21"
/codon_start=1
/product="fts21"
/protein_id="BAC57986.1"
/db_xref="GI:28804576"
/translation="MVQVRPSLTNSCHGSPVFASAPPAICSGVGAARCCALPS
LHGKVSLLSPYLRSGTLQHDCKRAEWQAKSAGTRIFASMPDMSARLKVIGVG
GGNNAINRMIGSLQGVFEFALNTDAQLQSAATHRVQIGETLRGLTGGNPELG
EKAABESLEIAEAVSDADLVFTAGMGGTGSGAPVVARLAKKEGGQITVGVTYPF
TFEGRRAOQGLEAIQLKKNVDLIVIPNDRLLDVVQATPLQAFLLADVLQGV
QGISDITITPGIIVVDFADVKAVMNSGTMLGVSGTCKNRAEAAQATSAFLIER
SIERATGVVNTTGGKDLTLOEVNRYSOVVTGLADPAANIIFGAVVDEKVTGAVHTI
IATGFSQTQKTLIDPKVARQEQDPSKGVDSFWKRPAPVSSRFQGLGSKGFL"

ORIGIN

Query Match 85.6%; Score 15.4; DB 8; Length 3926;
Best Local Similarity 94.1%; Pred. No. 2.1e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 AGAACACAGATTCCG 18

|||||

Db 3487 AGAACACAGATTCCG 3503

RESULT 50

AX463544

LOCUS

AX463544

DEFINITION

Sequence 108 from Patent WO0248337.

AC0911102

AX463544

VERSION AX463544.1 GI:21886318
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Duggan B.M., Yao M.G. and Griffin, J.A.
TITLE Secreted proteins
JOURNAL Patent: WO 0248337-A 108 20-JUN-2002;
INCYTE GENOMICS INC (US)
FEATURES
 Location/Qualifiers
 1. .4717
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"
 /note="Incyte ID No: 7503512CB1"
source
ORIGIN
 Query Match 85.6%; Score 15.4; DB 6; Length 4717;
 Best Local Similarity 94.1%; Pred. No. 2.1e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
OY 2 AGAACAAACAGATTCCG 18
 |||
Db 368 AGCACAAACAGATTCCG 384
Search completed: December 3, 2004, 03:06:23
Job time : 635.921 secs

This Page Blank (uspto)